AD 216837

Armed Services Technical Information Agency

ARLINGTON HALL STATION
ARLINGTON 12 VIRGINIA

FOR
MICRO-CARD
CONTROL ONLY

1 OF 2

UNCLASSIFIED

FC

UNIVERSITY OF MINNESOTA

SCHOOL OF
BUSINESS ADMINISTRATION

June 1, 1959

SOME PROPERTIES OF THE SPEARMAN ESTIMATOR IN BIOASSAY

Byron William Brown, Jr. 1/

Technical Report No. 6

University of Minnesota
Minneapolis, Minnesota

1/ Work done in part under contract Nonr2582(00), Task NR 042-200 of the Office of Naval Research.

TABLE OF CONTENTS

Page

GLOSSARY

1. The Tolerance Distribution

x	:	a dose level
$F(x)$	:	a tolerance distribution
$f(x)$	:	$F'(x)$
$f^{(n)}(x)$	:	the nth derivative of f
μ	:	the mean of the tolerance distribution
x_m	:	the mode of F
$x_{.50}$	:	the median of F
σ	:	the standard deviation of F
R	:	the distance between the 20th and 80th percentiles of F
f_m	:	$f(x_m)$, the maximum of $f(x)$.

2. The Experiments

x_i	:	the ith dose level
n_i	:	the number of subjects tested at x_i
n	:	the number of subjects tested at each x_i when the sample sizes are all equal
d	:	the common distance between dose levels: $d=x_{i+1}-x_i$
ρ	:	d/σ
ρ'	:	d/R
x_o	:	the "middle" dose level ($x_i=x_o+id$, $i=0,\pm1,\pm2,\ldots$)
k	:	the number of dose levels on each side of x_o for the finite experiment
a	:	the distance on each side of x_o covered by dose levels in the finite experiment: $a=kd$

N	:	the total number of subjects tested in the finite experiment: $N=n(2k+1)$

3. Observations and Estimators

r_i	:	the observed number of subjects responding at the ith dose level, x_i
p_i	:	the proportion of subjects responding at the ith dose level: $p_i=r_i/n_i$
$\bar{x}$	:	the Spearman estimator

4. Information and Characteristics of Estimators

E ( )	:	denotes expectation when a random variable appears in the brackets
V ( )	:	denotes variance
MSE ( )	:	denotes mean square error
B ( )	:	denotes bias
subscript a or k	:	denotes finite experiment (doses $x_i=x_o+id$, $i=0, \pm1, \pm2, \ldots \pm k$). If no subscript a or k appears, the experiment is infinite (doses $x_i=x_o+id$, $i=0, \pm1, \pm2, \ldots\ldots$). If no conditional notation concerning x_o appears, x_o is taken to be randomly chosen on the interval (o,d).
E_{x_o} ( )	:	expectation with respect to x_o over the interval (0,d)
subscript A	:	denotes an asymptotic moment (see (6.3) and (9.1))
$\bar{V}$	:	the value $\frac{d}{n}\int F(1-F)dx$. $\bar{V}$ is shown in section ...3 to approximate $V(\bar{x})$ and is defined in section 6.3 to be the asymptotic variance of $\bar{x}$ for the infinite experiment
I	:	information for scale parameter known
I^{11}	:	element in the inverse information matrix corresponding to μ when the scale parameter is unknown

E	:	denotes efficiency relative to I if no estimator follows in brackets. $E(x_o)$ denotes efficiency conditional on x_o.
E^{11}	:	denotes efficiency relative to I^{11}

1. INTRODUCTION

1.1 An Example of a Quantal Assay:

In certain experimental situations it is impossible to measure the variable on each experimental unit directly, but it is possible to fix[1] a number for each unit and then determine simply whether or not the experimental unit has a measurement greater than this number. This type of experimental situation is found in a variety of fields of biological investigation. The following example from hormone assay illustrates the nature of the problem(13).

Some estrogenic preparations are extracts from the urine of pregnant mares. These preparations are mixtures of several estrogens. The estrogenic strength of such extracts can not be measured analytically in a satisfactory way. It is known, however, that if sufficient estrogenic substance is given to immature or spayed female mice, they will show cornification of the vagina. If a fixed dose of the preparation is given to a test animal and cornification is observed, all that is known is that the dose administered was at least as great as the 'tolerance' of the mouse to this preparation. If no cornification is observed, there is no way of knowing how

[1]In some applications this number can be fixed only with appreciable error. This case is not considered in this paper, but has been discussed by Haley (21) for certain parametric formulations.

much higher the dose would have had to be to induce the response. It is desirable to estimate the mean tolerance of a group of test animals to the preparation on the basis of such data. The mean tolerance and the strength are inversely related. Other examples of this assay situation can be found in insecticide research (7), vitamin research (20), vaccine screening for safety (25), and toxicity evaluation of various chemicals (27). The quantal assay situation can also be found in industry in munitions testing (14) and reliability testing (30) among other applications.

1.2 Terminology

The variable under investigation will be called the dose. The dose may be a direct measure of the stimulus (e.g. concentration of an injection) or it may be some transformation of this direct measurement (commonly the log of the measurement). The experimental units will be tested at various doses. The observation on each unit will be either a response or a no-response. All-or-none responses are called quantal responses in biological experiments. The experiment is a quantal assay. The probability of response depends on the dose. The function relating the probability of response to the dose level is the tolerance distribution or dose response function.

1.3 The Experiment and Model

The usual quantal assay is done in the following manner. A set of dose levels is chosen. The test subjects to be used in the experiment are randomly allocated to the dose levels. The number of responses among the subjects in each dose level group is used as the basis for inference.

The test animals should be randomly selected from a well-defined population of subjects. The dose response function is descriptive of this population. The stability of the dose response function over time must be investigated (3).

The notation used in this paper will be as follows:

x denotes a dose level

$F(x)$ denotes the dose response function, i.e. the expected proportion of responses at dose x

x_i denotes the i^{th} dose level used in the experiment, $i=0, \pm 1, \pm 2, \ldots$

n_i denotes the number of subjects tested with dose x_i

r_i denotes the number of subjects responding among the n_i subjects receiving dose x_i

p_i denotes the proportion of subjects responding among the group of subjects receiving dose x_i, $p_i = {}^{r_i}/_{n_i}$

The assumptions concerning the dose response function and the observations are:

$F(x)$ is a distribution function

$F(x)$ has a first moment, μ, called the mean tolerance

The observations on the subjects are mutually independent

The observations are the dichotomous quantal response variables. A set of sufficient statistics for the experiment consists of the numbers of responses, r_i, among the subjects tested at each dose level. The r_i are mutually independently distributed binomial variables with means $n_i F(x_i)$, $i=0, \pm 1, \pm, \ldots$.

The primary problem in this paper is the estimation of μ on the basis of the experiment described above. Some biostatistical writers (4, 5, 6, 8, 10, 15, 16, 17, 18) recommend parametric estimators for this problem, i.e. estimators which necessitate the specification of a functional form for the dose response function. One function used frequently is the normal tolerance distribution:

$$F(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-t^2/2} \, dt \qquad (1.1)$$

This paper presents an evaluation of the Spearman estimator, a nonparametric estimator of μ.

2. THE SPEARMAN ESTIMATOR

2.1 Definition of the Spearman Estimator

The estimator to be discussed in this paper was described by Spearman (28) in 1908. He gives credit for the idea to the German psycho-physiologist, Muller. The estimator was described again by Karber (23) in 1931 and is occasionally referred to as the Spearman-Karber estimator.

Spearman defined the estimator for regularly spaced dose levels, $x_i = x_o + id$, $i=0, \pm 1, \pm 2, \ldots$, and equal numbers of subjects tested at a finite number of dose levels, say $n_i = n$ for $i=0, \pm 1, \pm 2, \ldots \pm k$. The estimator is

$$\bar{x} = \sum_{-k}^{k-1} (x_i + {}^d/2)\,(p_{i+1} - p_i).$$

$\bar{x}$ is analogous to a grouped mean for continuous data.

Spearman computed this estimator only when $p_{-k} = 0$ and $p_k = 1$. In practice, the following modification is used:

$$\bar{x} = p_{-k}(x_{-k} - {}^d/_2) + \sum_{-k}^{k-1} (x_i + {}^d/_2)\,(p_{i+1} - p_i) + (1-p_k)(x_k + {}^d/_2) \qquad (2.1)$$

Thus any estimate of probability, p_{-k}, below the lowest dose level (x_{-k}) for which $n_i \neq 0$ is assigned to the point half a dose interval below the lowest level; and the estimate of probability, $1-p_k$, above x_k is handled similarly.

Armitage and Allen (2) extended Spearman's definition to unequally spaced dose levels, x_i:

$$\bar{x} = \sum_{-k}^{k-1} \left(\frac{x_i + x_{i+1}}{2}\right) (p_{i+1} - p_i) \qquad (2.2)$$

This definition can be modified to allow for estimates of probability below x_{-k} and above x_k.

Irwin (22) and Finney (15, 16) discussed an experiment in which subjects would be tested at an infinity of dose levels. This experiment called for $n_i = n$ and $x_i = x_o + id$, $i = 0, \pm1, \pm2, \ldots.$ with x_o and d chosen arbitrarily. (See Appendix I for a discussion of the resulting infinite sample space.)

If $n_i \neq 0$ for all i and x_i are chosen so that $x_i \to \infty$ as $i \to \infty$ and $x_i \to -\infty$ as $i \to -\infty$, then the estimator can be defined as the limit of $\bar{x}$ defined in (2.2) as $k \to \infty$.

$$\bar{x} = \lim_{k \to \infty} \sum_{-k}^{k-1} \left[\frac{x_i + x_{i+1}}{2}\right] \left(p_{i+1} - p_i \right) \qquad (2.3)$$

The following experimental designs will be considered: (a) $x_i = x_o + id$ with $n_i = n$ for $i = 0, \pm 1, \pm 2, \ldots \pm k$ and $n_i = 0$ for $i = \pm(k+1), \pm(k+2), \ldots$. (b) $x_i = x_o + id$ with $n_i = n$ for $i = 0, \pm 1, \pm 2, \ldots$. The first experiment will be referred to as the finite experiment and the second experiment will be referred to as the infinite experiment.

Note: The Spearman estimator for the finite experiment (2.1) can be expressed in several ways:

$$\bar{x} = x_k + {}^{d}/_{2} - d \sum_{-k}^{k} p_i \qquad (2.4)$$

$$\bar{x} = x_o + {}^{d}/_{2} + d \sum_{1}^{k} q_i - d \sum_{-k}^{0} p_i \qquad (2.5)$$

$$q_i = 1 - p_i$$

2.2 The Mean and Variance of the Spearman Estimator for the Finite Experiment

The exact mean and variance of the Spearman estimator for the finite experiment and fixed x_o are:

$$E_k(\bar{x}|x_o) = (x_{-k} - {}^{d}/_{2})F_{-k} + \sum_{-k}^{k-1} (x_i + {}^{d}/_{2})(F_{i+1} - F_i) + (x_k + {}^{d}/_{2})(1 - F_k)$$

$$; \text{ where } F_i = F(x_i) \qquad (2.6)$$

$$V_k(\bar{x}|x_o) = \frac{d^2}{n} \sum_{-k}^{k} F_i(1-F_i) \qquad (2.7)$$

As an illustration (Table 2.1), the bias and the variance of the Spearman estimator have been computed for a normal tolerance distribution (1.1). The experimental design consists of five dose levels, two standard deviation units apart, with n subjects at each of the five levels. The bias and the variance of the estimator depend on the location of the dose mesh relative to the mean of the tolerance distribution. Therefore the bias and variance were computed for several locations of the dose mesh. The location of the dose mesh is indicated by the distance of the middle dose, x_0, from the mean of the tolerance distribution, μ, in standard deviation units.

When the mean is within the interval spanned by the dose levels, the fluctuations in the bias and the variance as functions of the location of the dose mesh are negligible. When the dose mesh fails to cover the mean the bias becomes large and the variance goes to zero. The mean square errors for the case of n equal to ten and to one hundred are also shown in Table 2.1.

2.3 Comparison of the Spearman Estimator with Parametric Competitors

It is clear from the above introduction to the Spearman estimator that it has certain advantages over its parametric competitors:

a) The Spearman estimator is simple in concept, being just the mean of a histogram reconstructed from the quantal data.

b) The Spearman estimator is simple to compute. It involves only the sum of the observed proportions (2.4). The

Tablo 2.1

The Bias, Variance and Mean Square Error of the Spearman Estimator (2.1) of the Mean of a Normal Tolerance Distribution, Using Five Dose Levels Spaced 2σ Apart

$\frac{x_o-\mu}{\sigma}$	$\frac{\text{Bias}}{\sigma}$	$\frac{nV_k(\bar{x}\mid x_o)}{\sigma^2}$	$nMSE_k(\bar{x}\mid x_o)/\sigma^2$ n=10	$nMSE_k(\bar{x}\mid x_o)/\sigma^2$ n=100
0	0	1.178	1.178	1.178
.2	-.003	1.168	1.168	1.169
.4	-.004	1.144	1.144	1.146
.6	-.005	1.114	1.114	1.117
.8	-.003	1.089	1.089	1.090
1.0	0	1.078	1.078	1.078
1.2	.003	1.089	1.089	1.090
1.4	.005	1.114	1.114	1.117
1.6	.005	1.144	1.144	1.147
1.8	.003	1.168	1.168	1.169
2.0	0	1.178	1.178	1.178
2.2	-.003	1.168	1.168	1.169
2.4	-.004	1.143	1.143	1.145
2.6	-.004	1.112	1.112	1.114
2.8	-.001	1.086	1.086	1.086
3.0	.003	1.073	1.073	1.074
3.2	.008	1.078	1.079	1.084
3.4	.014	1.095	1.097	1.115
3.6	.021	1.112	1.116	1.156
3.8	.031	1.114	1.124	1.210
4.0	.046	1.089	1.110	1.301
4.2	.069	1.030	1.078	1.506
4.4	.106	.936	1.048	2.060
4.6	.158	.815	1.065	3.311
4.8	.229	.678	1.202	5.922
5.0	.320	.539	1.563	10.779
5.2	.432	.410	2.276	
5.4	.563	.299	3.469	
5.6	.710	.208	5.249	
5.8	.872	.139	7.743	
6.0	1.046	.089	11.030	
6.2	1.228	.055		
6.4	1.416	.032		
6.6	1.609	.019		

parametric estimation procedures ordinarily involve either an iterative solution or a **weighted regression solution.**

c) The exact mean and variance of the Spearman estimator are easily obtained for any size sample. Therefore experimental design investigations are readily done. In contrast, only asymptotic theory is available for the parametric estimators.

d) The Spearman estimator is nonparametric in that no functional form need be assigned F(x) in order to compute the value of the estimate from the data. How important this point is depends on the robustness of the parametric estimators. Some relevant results are presented later in this paper (section 9).

2.4 Examples of Tolerance Distributions

The tolerance distributions in Table 2.2 serve as illustrations throughout this paper. The first four distributions are used in **practice.** The remaining distributions are to illustrate specific points (see sections 4.2, 4.3, 5.4, 7.3, and 7.6).

3. THE INFINITE EXPERIMENT: THE MEAN AND VARIANCE OF THE SPEARMAN ESTIMATOR

3.1 The Infinite Experiment

Irwin's (22) and Finney's (15, 16) concept of an infinite experiment (section 2.1) makes possible a mathematical discussion of the effect of the location of the dose mesh on

Table 2.2

Tolerance Distributions

NAME	FUNCTIONAL FORM		VARIANCE
1. Normal	$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\beta(x-\mu)} e^{-t^2/2}\,dt$;	$-\infty<x<\infty$	$\frac{1}{\beta^2}$
2. Logistic	$[1+e^{-\beta(x-\mu)}]^{-1}$;	$-\infty<x<\infty$	$\frac{\pi^2}{3\beta^2}$
3. Angular	$\sin^2[\beta(x-\mu)+\pi/4]$;	$-\pi/4\leq\beta(x-\mu)\leq\pi/4$	$\frac{\pi^2-8}{16\beta^2}$
4. One-Particle	$1-e^{-x/\mu}$;	$x>0$	μ^2
5. Uniform	$\beta(x-\mu)+1/2$;	$-1/2\leq\beta(x-\mu)\leq 1/2$	$\frac{1}{12\beta^2}$
6. Algebraic	$1-x^{-s}$;	$x\geq 1,\ s>1$	$\frac{s}{(s-2)(s-1)^2}$
7. Student's	$K_e\int_{-\infty}^{(x-\mu)} \frac{dt}{(1+\frac{t^2}{1+2e})^{1+e}}$;	$-\infty<x<\infty,\ e>0$	$\frac{2e+1}{2e-1}$

Note 1: For the algebraic distribution the mean is $\frac{s}{s-1}$.
For the other distributions the mean is μ.

Note 2: For the first five distributions β must be positive and the variance exists for all β. For the last two distributions the variance exists only if $s>2$ and $e>\frac{1}{2}$ respectively.

the bias and the variance of the estimator without discussing the possibilities of grossly misplacing the whole set of dose levels relative to μ. The concept of the infinite experiment also facilitates the development of large sample definitions of mean square error and efficiency in later sections (6 and 7). The investigation of the infinite experiment has practical importance since it is shown in Appendix I that the information for the infinite experiment is essentially the same as that for the corresponding finite experiment covering "most" of the range of F(x).

3.2 Mean and Variance of the Spearman Estimator for the Infinite Experiment

Lemmas 3.2.1, 3.2.2, and 3.2.3 establish conditions under which the Spearman estimator has the following mean and variance:

$$E(\bar{x}|x_0) = \sum_{-\infty}^{\infty}(x_i+{}^d/_2)(F_{i+1}-F_i) \qquad (3.1)$$

$$V(\bar{x}|x_0) = \frac{d^2}{n}\sum_{-\infty}^{\infty}F_i(1-F_i) \qquad (3.2)$$

Using (2.5), the estimator is:

$$\bar{x} = \lim_{k\to\infty}\left\{x_0+{}^d/_2+d\sum_{1}^{k}q_i-d\sum_{-k}^{0}p_i\right\} \qquad (3.3)$$

Similarly

$$\sum_{-\infty}^{\infty}(x_i+{}^d/_2)(F_{i+1}-F_i) = x_0+{}^d/_2+d\sum_{1}^{\infty}(1-F_i)-d\sum_{-\infty}^{0}F_i \qquad (3.4)$$

Lemma 3.2.1: If F has a first moment, μ, then the series

$$x_0+{}^d/_2+d\sum_{1}^{\infty}(1-F_i)-d\sum_{-\infty}^{0}F_i \qquad (3.5)$$

converges to a finite value.

Proof: The lemma is established if $d\sum_{1}^{\infty}(1-F_i)$ and $d\sum_{-\infty}^{0}F_i$ are shown to converge to finite values. Consider the remainder for $d\sum_{1}^{\infty}(1-F_i)$:

$$d\sum_{k+1}^{\infty}(1-F_i) \leq \int_{x_k}^{\infty}(1-F)dx$$

Interchanging the order of integration on the right:

$$d\sum_{k+1}^{\infty}(1-F_i) \leq \int_{x_k}^{\infty} tdF(t)$$

If F has a first moment, the integral on the right goes to zero as x_k becomes infinite. Therefore $d\sum_{1}^{\infty}(1-F_i)$ is finite. Similarly $d\sum_{-\infty}^{0}F_i$ can be shown to be finite. Q.E.D.

Lemma 3.2.2: If F has a first moment, then the series

$$\frac{d^2}{n}\sum_{-\infty}^{\infty}F_i(1-F_i) \qquad (3.6)$$

converges.

Proof: In the proof of Lemma 3.2.1 it was established that $\sum_{1}^{\infty}(1-F_i)$ is finite. Since $0 \leq F_i(1-F_i) \leq 1-F_i$ the series $\sum_{1}^{\infty}F_i(1-F_i)$ also is finite. Similarly $\sum_{-\infty}^{0}F_i(1-F_i)$ is finite, and, therefore, $\frac{d^2}{n}\sum_{-\infty}^{\infty}F_i(1-F_i)$ is finite. Q.E.D.

Lemma 3.2.3: If F has a first moment, the Spearman estimator (3.3) for the infinite experiment converges with probability one to a random variable with mean and variance given by (3.1) and (3.2).

Proof: Lemmas 3.2.1 and 3.2.2 prove the convergence of the series in (3.5) and (3.6). Theorem 2.3 in chapter III of

Doob (12) establishes the sufficiency of the convergence of these two series for the convergence of $\bar{x}$ with probability one, and shows that the expected value and variance of $\bar{x}$ are given respectively by (3.5) and (3.6), or, using (3.4), by (3.1) and (3.2). Q.E.D.

The variance of the Spearman estimator is not zero even though the number of observations is infinite. Similarly, it is shown in Appendix I that the information contained in the infinite experiment is finite in the common parametric formulations.

EFFECT OF DOSE MESH LOCATION ON BIAS AND VARIANCE OF THE SPEARMAN ESTIMATOR IN THE INFINITE EXPERIMENT

1 General Discussion

In the infinite experiment the expected value and variance of the estimator depend, in general, on the doses, $x_i=x_o+id$, $i=0,\pm1,\pm2,\ldots$. For a particular spacing, d, the expected value and variance will be simply functions of x_o with period d.

Finney (15,16) has computed the bias and variance for normal and logistic tolerance distributions. Since the Spearman estimator is nonparametric, it is desirable to have information on the bias over a wide class of distribution functions. It is possible to find bounds for the bias, distributions that maximize the bias, and conditions on F(x) that limit the bias. Bounds for the fluctuation in the variance can also be obtained.

.2 Bounds on the Bias of the Spearman Estimator

4.2.1 Expression for the Bias

The bias of an estimator is the difference between its expected value and the parameter estimated. Denote the bias for the Spearman estimator for the infinite experiment, conditional on x_o, by $B(\bar{x}|x_o)$. Using (3.1):

$$\begin{aligned} B(\bar{x}|x_o) &= E(\bar{x}|x_o)-\mu \\ &= \sum_{-\infty}^{\infty}(x_i+{}^d/_2)(F_{i+1}-F_i)-\int_{-\infty}^{\infty} x dF(x) \\ &= \sum_{-\infty}^{\infty}(x_i+{}^d/_2)(F_{i+1}-F_i)-\sum_{-\infty}^{\infty}\int_{x_i}^{x_{i+1}} x dF(x) \\ &= \sum_{-\infty}^{\infty}(x_i+{}^d/_2-c_i)(F_{i+1}-F_i) \end{aligned} \tag{4.1}$$

where
$$c_i = \frac{\int_{x_i}^{x_{i+1}} x dF(x)}{\int_{x_i}^{x_{i+1}} dF(x)}. \tag{4.2}$$

4.2.2. Bound on the Bias

From (4.1) it follows that:

Lemma 4.2.2

$$|B(\bar{x}|x_o)| \leq \sum_{-\infty}^{\infty}|x_i+{}^d/_2-c_i|(F_{i+1}-F_i) \leq \frac{d}{2}\sum_{-\infty}^{\infty}(F_{i+1}-F_i)=\frac{d}{2} \tag{4.3}$$

4.2.3. Tolerance Distributions for which the Bound is Attained

The bound given in (4.3) is attained, e.g. the one point distribution. In this case when one of the dose levels coincides with the mass point of the distribution, the true mean of the distribution equals the dose level, but the estimator has an expected value $^d/_2$ units below this dose level (at the midpoint of the dose interval showing the probability increment). This example indicates that for any distribution function, if d is large relative to the dispersion of $F(x)$, the bias can be approximately $d/2$.

Even if one excludes the one point distribution a bias of $d/2$ is attainable. Consider the class of discrete distributions with mass points on a lattice with spacing D. Then if the dose mesh has $d=D/m$, m a positive integer, and if the dose mesh is located so that the mass points coincide with dose levels, then each c_i (4.2) has the value x_i and

$$B(\tilde{x}|x_o) = \frac{d}{2}\sum_{-\infty}^{\infty}(F_{i+1}-F_i) = \frac{d}{2}$$

4.2.4. Bound on the Bias for Unimodal Distributions

The situations discussed in paragraph 4.2.3 do not occur often in practice. In this paragraph $F(x)$ is restricted to functions with the usual properties possessed by tolerance distributions.

Lemma 4.2.4. If $F(x)$ has a unimodal density with maximum ordinate, f_m, then

$$\left|B(\tilde{x}|x_o)\right| \leq \frac{d^2}{8} f_m \qquad (4.4)$$

(Unimodality means the density, $f(x)$, is non-decreasing for x less than the mode, x_m, and non-increasing for x greater than the mode.)

Proof: The bound can be obtained by examining the terms in the expression for the bias (4.1).

Let $I_i=(x_i,x_{i+1})$ be an interval located above the modal value x_m, so that $f(x)$ is non-increasing in I_i. Then, using c_i as defined in (4.2),

$c_i \leqq x_i + {}^d/_2$ no matter what the density $f(x)$ is in I_i.

For the given probability $F_{i+1}-F_i$ associated with I_i the minimum value possible for c_i is attained for the density, g:

$$g(x) = f_i \qquad x_i \leqq x \leqq x_i + Rd$$

$$g(x) = f_{i+1} \qquad x_i + Rd < x < x_{i+1}$$

R is determined by:

$$\int_{x_i}^{x_{i+1}} g(x)dx = F_{i+1}-F_i \quad ,$$

$$\text{i.e.} \quad Rd(f_i - f_{i+1}) + df_{i+1} = F_{i+1} - F_i \quad .$$

For the minimizing density, $g(x)$, the value of c_i becomes:

$$c_{ig} = \frac{(x_i + \frac{Rd}{2})Rd(f_i - f_{i+1}) + (x_i + {}^d/_2)f_{i+1}d}{F_{i+1}-F_i}$$

Since c_i for the density $f(x)$ is bounded below by c_{ig} and above by $x_i + {}^d/_2$, the bias term for I_i satisfies

$$0 \leqq (x_i + {}^d/_2 - c_i)(F_{i+1}-F_i) \leqq (x_i + {}^d/_2 - c_{ig})(F_{i+1}-F_i)$$

The term on the right can be evaluated in terms of d, R, and the values of $f(x)$ to obtain:

$$0 \leqq (x_i + {}^d/_2 - c_i)(F_{i+1}-F_i) \leqq \frac{d^2}{2} R(1-R)(f_i - f_{i+1})$$

Since $0 \leqq R \leqq 1$, this inequality can be relaxed to obtain:

$$0 \leqq (x_i + {}^d/_2 - c_i)(F_{i+1}-F_i) \leqq \frac{d^2}{8} (f_i - f_{i+1}) \qquad (4.6)$$

Similarly, for intervals below the mode, the following inequality is obtained:

$$-\frac{d^2}{8}(f_{i+1}-f_i) \leq (x_{i+1}+{}^d/_2-c_i)(F_{i+1}-F_i) \leq 0 \qquad (4.7)$$

Also, for the interval containing the mode, say (x_o,x_1), the inequality obtained is:

$$-\frac{d^2}{8}(f_m-f_o) \leq (x_o+{}^d/_2-c_o)(F_1-F_o) \leq \frac{d^2}{8}(f_m-f_1) \qquad (4.8)$$

Then (4.6), (4.7), and (4.8) combine to give a bound on the sum of the contributions to the bias from all of the intervals:

$$|B(\bar{x}|x_o)| \leq \max \begin{cases} \frac{d^2}{8}(f_m-f_1) + \sum_1^\infty \frac{d^2}{8}(f_i-f_{i+1}) \\ \\ \frac{d^2}{8}(f_m-f_o) + \sum_{-\infty}^0 \frac{d^2}{8}(f_{i+1}-f_i) \end{cases} \qquad (4.9)$$

Then, since $\sum_1^\infty (f_i-f_{i+1})=f_1$ and $\sum_{-\infty}^0 (f_{i+1}-f_i)=f_o$, the bound is

$$|B(\bar{x}|x_o)| \leq \frac{d^2}{8} f_m$$

Q. E. D.

4.2.5. Unimodal Tolerance Distributions for Which the Bound is Attained

The bound (4.4) is attained for certain unimodal densities. Consider the following example: Let $f(x)=1$ for $0\leq x\leq 1$, $f(x)=0$ otherwise. Let the dose interval, d, be given by:

$$d = \frac{1}{N+{}^1/_2} \qquad \text{(N a positive integer)}$$

Let $x_o=0$. The the dose levels will be given by:

$$\cdots \frac{-1}{N+{}^1/_2}, \frac{0}{N+{}^1/_2}, \frac{1}{N+{}^1/_2}, \cdots \frac{N}{N+{}^1/_2}, \frac{N+1}{N+{}^1/_2}, \cdots$$

for all intervals except the one containing the point, $x=1$, $f(x)$ is uniform and $c_i = x_i + {}^d/_2$ so that the contributions to the bias are zero. For the interval containing $x=1$, i.e. the interval

$$\left(\frac{N}{N+{}^1/_2}, \frac{N+1}{N+{}^1/_2} \right),$$

the contribution to the bias can be computed as follows:

$$c_N = \frac{N}{N+{}^1/_2} + \frac{1}{2}\left(\frac{N+1}{N+{}^1/_2} - \frac{N}{N+{}^1/_2}\right) = \frac{N+{}^1/_4}{N+{}^1/_2}$$

$$x_N + {}^d/_2 = 1$$

$$F_{N+1} - F_N = \frac{1}{2(N+{}^1/_2)}$$

$$B(\bar{x}|x_o) = \left[\frac{N+{}^1/_4}{N+{}^1/_2} - 1 \right] \frac{1}{2(N+{}^1/_2)} = \frac{1}{(N+{}^1/_2)^2} \quad \frac{1}{8}$$

This is the bound given in (4.4) since $d = \frac{1}{N+{}^1/_2}$ and $f_m = 1$.

The uniform distribution has the properties of symmetry and only two points of inflection. Therefore, these properties do not lead to a stronger bound on the bias of the Spearman estimator.

4.2.6. Bounds on the Bias for Distributions in Terms of Derivatives

The example of the uniform distribution in paragraph 4.2.5 suggests the contributions to the bias of the Spearman estimator come from discontinuities and rapid rates of change in the density function. Bounds on the bias can be tightened with bounds on $f'(x)$ or higher derivatives.

Lemma 4.2.6 If $F(x)$ is differentiable s times, $F(x)$ is symmetrical, and $f^{(n)}(x)$ has limit zero for $x \to +\infty$ and for $x \to -\infty$, $n=0,1,2,\dots s$, then

$$|B(\bar{x}|x_o)| \leq \sup_x |P_{n+1}(x)| d^{n+1} \int_{-\infty}^{\infty} |f^{(n)}(x)| dx \qquad (4.10)$$

$n=0,1,2,\dots s$

$f^{(n)}(x)$ is the n^{th} derivative of $f(x)$

(See Appendix II for definition of $P_n(x)$--the nth Bernoulli function).

Proof: Let $\bar{x}_k$ be the Spearman estimator for a finite number of dose levels, $x_i=x_o+id$, $i=0,\pm1,\pm2,\dots,\pm k$. From (2.5):

$$\bar{x}_k = x_o - d\sum_{-k}^{k}(p_i - 1/2)$$

$$E(\bar{x}_k|x_o) = x_o - d\sum_{-k}^{k}(F(x_o+id) - 1/2)$$

Then by the Euler-MacLaurin formula (see Appendix II):

$$E(\bar{x}_k|x_o) = x_o - d\Bigg[\int_{-k}^{k}\left(F(x_o+xd) - 1/2\right)dx + 1/2\left(F(x_o+kd) - 1/2\right) + \frac{1}{2}\left(F(x_o-kd) - 1/2\right) - d\int_{-k}^{k} P_1(x) f(x_o+xd)dx\Bigg] \qquad (4.11)$$

From (2.6) and (.1) it is seen that $E(\bar{x}|x_o) = \lim_{k \to \infty} E(\bar{x}_k|x_o)$.

First consider the limit of the middle two terms within the brackets, i.e.,

$$\frac{1}{2}\left[F(x_o+kd)-1+F(x_o-kd)\right] \quad .$$

Since F is a distribution function, the limit of this expression as k becomes large is zero.

The first integral in (4.11) can be rewritten:

$$\int_{-k}^{k}\left[F(x_o+xd)-{}^1/_2\right]dx = \frac{1}{d}\int_{\mu-kd+(x_o-\mu)}^{\mu+kd+(x_o-\mu)}\left[F(y)-{}^1/_2\right]dy$$

$$= \frac{1}{d}\int_{\mu-(kd-(x_o-\mu))}^{\mu+(kd-(x_o-\mu))}\left[F(y)-{}^1/_2\right]dy + \frac{1}{d}\int_{\mu+kd-(x_o-\mu)}^{\mu+kd+(x_o-\mu)}\left[F(y)-{}^1/_2\right]dy$$

$$= \frac{1}{d}\int_{\mu+kd-(x_o-\mu)}^{\mu+kd+(x_o-\mu)}\left[F(y)-{}^1/_2\right]dy$$

(Since F(y) is assumed symmetric, $F(y)-\frac{1}{2}$ will be an odd function with respect to $y=\mu$.) Consider the limit of this integral as k becomes large. The length of the interval of integration remains constant. The value of the integrand approaches $\frac{1}{2}$. Hence,

$$\lim_{k\to\infty}\frac{1}{d}\int_{\mu+kd-(x_o-\mu)}^{\mu+kd+(x_o-\mu)}\left[F(y)-{}^1/_2\right]dy = \frac{x_o-\mu}{d}$$

Substituting the limiting values obtained thus far in (4.11), the expression for $E(\bar{x}|x_o)$ becomes:

$$E(\bar{x}|x_o) = \lim_{k\to\infty} E(\bar{x}_k|x_o) = \mu + \lim_{k\to\infty} d^2 \int_{-k}^{k} P_1(x) f(x_o + xd)\,dx$$

Thus the bias can be bounded as follows:

$$\left|B(\bar{x}|x_o)\right| \leq \sup_x \left| P_1(x) \right| d \int_{-\infty}^{\infty} \left| f(x) \right| dx$$

Before taking limits the integral involving P_1 and f could be integrated by parts, making use of the relationship:

$$P_n'(x) = (-1)^{n-1} P_{n-1}(x)$$

Repeated integration by parts would lead to the general expression of (4.10). Q.E.D.

4.2.7. Expressions for the Bound Involving Derivatives of the Tolerance Distribution

When the bound (4.10) is evaluated for n=0 and n=1, the results are respectively:

$$|B(\bar{x})| \leq d/2$$

$$|B(\bar{x})| \leq \frac{d^2}{6} f_m$$

The first bound is identical with the one obtained in paragraph 4.2.2. without the assumption of symmetry. The second bound is of the same order in d as that obtained in paragraph 4.2.4 but the constant $1/6$ is greater than $1/8$.

Lemma 4.2.6 provides a sequence of bounds on the bias in increasing powers of d. For example, if n is taken to be 2, the expression for the bound becomes:

$$\left|B(\bar{x}|x_o)\right| = \left[\sup_x \left| \sum_1^\infty \frac{\sin k\, 2\pi\, x}{2^2 \pi^3\, k^3} \right|\right] d^3 \int_{-\infty}^{\infty} \left| f''(x) \right| dx$$

From Appendix II:

$$\sup_x \left| \sum_1^\infty \frac{\sin k 2\pi x}{2^2 \pi^3 k^3} \right| = .0080\ldots$$

A simple expression for the integral can be obtained if the density $f(x)$ is assumed to have exactly two points of inflection, $x=\mu \pm c$, with the absolute value of the derivative of $f(x)$ at these two points being f_c'. Then

$$\int_{-\infty}^{\infty} \left| f''(x) \right| dx = \int_{-\infty}^{\mu-c} f''(x)dx - \int_{\mu-c}^{\mu+c} f''(x)dx + \int_{\mu+c}^{\infty} f''(x)dx$$

$$= 4f_c'$$

Thus the bound for the bias, using n=2, becomes:

$$\left|B(\bar{x}|x_o)\right| \leq .032 d^3 f_c' \qquad (4.12)$$

For a symmetrical density function with two points of inflection, maximum slope f_c' and maximum ordinate f_m:

$$f_m \leq \sqrt{f_c'}$$

Thus, for this class of densities, the bounds given in (4.4) and (4.12) can be combined to g. ..

Theorem 4.2.7

$$|B(\bar{x}|x_0)| \leq \min.(.125d^2\sqrt{f_c'}\ ,\ .032d^3 f_c') \qquad (4.12a)$$

4.2.8 Computation of the Bounds on the Bias for Some Tolerance Distributions

If d is expressed in units of the standard deviation or some interpercentile difference of the tolerance distribution, the magnitudes of the various bounds on the bias obtained above can be more readily compared. Bound (4.3) becomes:

$$|B(\bar{x}|x_0)| \leq \frac{\rho}{2}\ \sigma \qquad (\rho = \frac{d}{\sigma}) .$$

Thus to assure that the bias is less than, say, 10 percent of the standard deviation of the tolerance distribution, d should be less than 20 percent of σ.

If the bound (4.4) is to be used, the modal ordinate must be specified. In Table 4.1 bounds on the bias, computed from (4.4), are given both in terms of σ and R. R is the distance from the 20th to the 80th percentile. The bounds on the bias in terms of σ and R for the various distributions are quite similar. For the four distributions, other than the special one-particle function, a choice of $d \leq 1.3\sigma$ will assure a bias of less than 10 percent of σ.

The bound given in (4.12) can be applied to the first three of the tolerance distributions listed in Table 4.1.

Table 4.2 gives the bounds for the three tolerance distributions in terms of both σ and R. Note from Table 4.2 that if $d \leqq 2.8\sigma$ the bias of the Spearman estimator will be less than 10 percent of σ (using 4.12).

4.3 Effect of Dose Mesh Location on the Variance of the Spearman Estimator

4.3.1. Bound on $|V(\bar{x}|x_o)-\bar{V}|$

Let $x_{.50}$ be the median of F(x) (or a median if $F(x)=\frac{1}{2}$ does not uniquely determine $x_{.50}$). Then $F(t)\left(1-F(t)\right)$ is non-decreasing for $x \leqq x_{.50}$ and is non-increasing for $x \geqq x_{.50}$. Number the dose intervals so that $x_o \leqq x_{.50} < x_1$. Then

$$\int_{x_{i-1}}^{x_i} F(t)\left(1-F(t)\right) dt \leqq dF_i\left(1-F_i\right) \quad i=0,-1,-2,\ldots$$

$$\int_{x_i}^{x_{i+1}} F(t)\left(1-F(t)\right) dt \leqq dF_i\left(1-F_i\right) \quad i=1,2,\ldots$$

$$\int_{x_o}^{x_i} F(t)\left(1-F(t)\right) dt \leqq d\cdot\frac{1}{2}\cdot\frac{1}{2}$$

Combining these inequalities:

$$d\sum_{-\infty}^{\infty} F(x_i)\left[1-F(x_i)\right] + d/4 \geqq \int_{-\infty}^{\infty} F(t)\left[1-F(t)\right] dt$$

Table 4.1

Bounds for the Bias of the Spearman Estimator

for the Infinite Experiment 1/

Tolerance Distribution	Bound for the Bias 1/ as a Proportion of σ 2/	as a Proportion of R 3/
Logistic	$.0567\, \rho^2 \sigma$ 2/	$.0866\, \rho'^2 R$ 3/
Normal	$.0499\, \rho^2 \sigma$	$.0839\, \rho'^2 R$
Angular	$.0427\, \rho^2 \sigma$	$.0804\, \rho'^2 R$
Uniform	$.0361\, \rho^2 \sigma$	$.0750\, \rho'^2 R$
One-Particle	$.1250\, \rho^2 \sigma$	$.1732\, \rho'^2 R$

1/ using (4.4) for the bound, $\left|B(\bar{x}|x_o)\right| \leq \frac{d^2}{8} f_m$

2/ σ is the standard deviation of the tolerance distribution and $d = \rho\, \sigma$

3/ R is the distance between the 20^{th} and the 80^{th} percentile and $d = \rho'\, R$.

Table 4.2

Bounds for the Bias of the Spearman Estimator for the Infinite Experiment [1]

Tolerance Distribution	Bound on the Bias [1] As a Proportion of σ [2]	As a Proportion of R [3]
Logistic	$.0101\, \rho^3\sigma$ [2]	$.0568\, \rho'^3$ R [3]
Normal	$.0077\, \rho^3\sigma$	$.0529\, \rho'^3$ R
Angular	$.0075\, \rho^3\sigma$	$.0639\, \rho'^3$ R

[1] using (4.12) for the bound, $|B(\bar{x}\,|\,x_o)| \leq .032d^3 f'_m$

[2] σ is the standard deviation of the tolerance distribution and $d=\rho\sigma$

[3] R is the distance between the 20^{th} and 80^{th} percentile and $d=\rho' R$.

Note: Finney (15, 16) gives the actual maximum bias for the normal and logistic distributions for various values of d. For ρ of 2, $B=.005\sigma$ for the normal distribution while the table gives $.062\sigma$ as the bound; at $\rho=3$ Finney has $B=.107$ compared with a bound of $.208\sigma$ in the table.

Therefore the variance (3.2) satisfies the inequality:

$$V(\bar{x}|x_o) = \frac{d^2}{n} \sum_{-\infty}^{\infty} F(x_i)\left[1-F(x_i)\right] \geq \frac{d}{n}\int_{-\infty}^{\infty} F(t)\left[1-F(t)\right]dt - \frac{d^2}{4n}$$

Similarly an upper bound for the variance of the estimator can be obtained:

$$V(\bar{x}|x_o) \leq \frac{d}{n}\int_{-\infty}^{\infty} F(t)\left[1-F(t)\right] dt + \frac{d^2}{4n}$$

Let $\bar{V}$ be defined as:

$$\bar{V} = \frac{d}{n}\int_{-\infty}^{\infty} F(t)\left[1-F(t)\right] dt \qquad (4.13)$$

Then

$$\left|V(\bar{x}|x_o) - \bar{V}\right| \leq \frac{d^2}{4n} \qquad (4.14)$$

4.3.2. <u>A Tolerance Distribution for Which $\left|V(\bar{x}|x_o)-\bar{V}\right|$ Approximates the Bound Arbitrarily Closely</u>

The bound (4.14) on the deviation of $V(\bar{x}|x_o)$ from $\bar{V}$ is the supremum. Consider the two point distribution with masses of $\frac{1}{2}$ at the points 0 and 1. Then F(x) has the form:

$$F(x) = 0 \qquad x < 0$$
$$F(x) = 1/2 \qquad 0 \leq x < 1$$
$$F(x) = 1 \qquad 1 \leq x$$

Then

$$\bar{V} = \frac{d}{n}\int_{-\infty}^{\infty} F(t)\left[1-F(t)\right] dt = \frac{d}{4n} = \frac{d^2}{4n} \cdot \frac{1}{d} \quad .$$

If $d = 1-e$, $1 > e > 0$, then it is possible for two doses to lie between 0 and 1. In this case,

$$V(\bar{x}|x_o) = \frac{d^2}{n}\left(\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}\right) = \frac{d^2}{2n}$$

$$V(\bar{x}|x_o)-\bar{V} = \frac{d^2}{4n}\cdot 2-\frac{d^2}{4n}\cdot\frac{1}{d}$$

$$V(\bar{x}|x_o)-\bar{V} = \frac{d^2}{4n}\left(2-\frac{1}{1-e}\right)$$

Therefore, as e goes to zero the deviation of $V(\bar{x}|x_o)$ from $\bar{V}$ can be arbitrarily close to $+\frac{d^2}{4n}$.

Also, using the same example, but choosing $d = 1+e$,

$$V(\bar{x}|x_o)-\bar{V} = -\frac{d}{4n}\cdot\frac{1}{1+e}$$

if no dose levels lie between 0 and 1, so that the deviation can be arbitrarily close to $-\frac{d^2}{4n}$.

4.3.3. Bounds on $\left|V(\bar{x}|x_o)-\bar{V}\right|$ in Terms of Derivatives

The Euler-MacLaurin formulae yield better bounds on the fluctuation of $V(\bar{x}|x_o)$ due to the placement of x_o, when more stringent conditions are imposed on $F(x)$. If $F(x)$ has a continuous density (and the first moment of F exists, as has already been assumed) then (see Appendix II):

$$\frac{d^2}{n}\sum_{-\infty}^{\infty} F(x_o+id)\left[1-F(x_o+id)\right] = \frac{d^2}{n}\int_{-\infty}^{\infty} F(x_o+xd)\left[1-F(x_o+xd)\right]dx$$

$$-\frac{d^3}{n}\int_{-\infty}^{\infty} P_1(x)f(x_o+xd)\left[1-2F(x_o+xd)\right]dx$$

or

$$\left|V(\bar{x}|x_o)-\bar{V}\right| \leq \left|\frac{d^3}{n}\int_{-\infty}^{\infty} P_1(x)f(x_o+xd)\left[1-2F(x_o+xd)\right]dx\right| \quad (4.15)$$

Integrating by parts in (4.15), which is justified if F' is continuous and if the resulting integral exists:

$$|V(\bar{x}|x_o)-\bar{V}| \leq \left| \frac{d^4}{n} \int_{-\infty}^{\infty} P_2(x) \left[\frac{d}{dx} f(x_o+xd)\left(1-2F(x_o+xd)\right)\right] dx \right|$$

First assume that F(1-F) has exactly two points of inflection, say at $x=c_1$ and at $x=c_2$. Then the integrand is positive for $x < c_1$ and for $x > c_2$ and negative for x between c_1 and c_2. Then

$$|V(\bar{x}|x_o)-\bar{V}| \leq \frac{d^3}{6n}\left\{ f(c_1)\left[1-2F(c_1)\right] - f(c_2)\left[1-2F(c_2)\right]\right\}$$

If F is symmetrical, then:

$$|V(\bar{x}|x_o)-\bar{V}| \leq \frac{d^3}{3n} f(c_1)\left[1-2F(c_1)\right]$$

Also

$$|V(\bar{x}|x_o)-\bar{V}| \leq \frac{d^3}{3n} \cdot f_m \qquad (4.16)$$

4.3.4 <u>Computation of Bounds on $|V(\bar{x}|x_o)-\bar{V}|$ for Some Tolerance Distributions</u>

To illustrate the magnitude of the bounds on $|V(\bar{x}|x_o)-\bar{V}|$ express d in units of the standard deviation or an interpercentile deviation of F(x) and express $|V(\bar{x}|x_o)-\bar{V}|$ as a proportion of $\bar{V}$. Thus, using (4.14) and (4.16):

$$\frac{|V(\bar{x}|x_o)-\bar{V}|}{\bar{V}} = \frac{d/4}{\int_{-\infty}^{\infty} F(x)\left[1-F(x)\right]dx} \qquad (4.14a)$$

$$\frac{|V(\bar{x}|x_o)-\bar{V}|}{\bar{V}} = \frac{\frac{d^2}{3} f_m}{\int_{-\infty}^{\infty} F(x)\left[1-F(x)\right]dx} \qquad (4.16a)$$

See Table 4.3 for numerical examples. From the second column, if d is less than $.4\sigma$ then $\frac{|V(\bar{x}|x_o)-\bar{V}|}{\bar{V}} \leq$ 20 percent for each of the five tolerance distributions. The fourth column indicates that for the logistic, normal and angular distributions, a d less than $.8\sigma$ assures that $\frac{|V(\bar{x}|x_o)-\bar{V}|}{\bar{V}} \leq$ 20 percent.

5. RANDOM LOCATION OF THE DOSE MESH IN THE INFINITE EXPERIMENT

5.1 Introduction

Random location of the dose mesh is accomplished for the infinite experiment by fixing the dose interval d and randomly choosing the dose level, x_o, from the uniform distribution over the interval (0,d). Even though no effort is made to randomly locate the dose mesh, in certain routine screening procedures, at least, the tolerance distributions are essentially randomly located with respect to the fixed dose mesh.

5.2 Unbiasedness of the Spearman Estimator

The expected value and variance of the Spearman estimator for random choice of x_o will be denoted by $E(\bar{x})$ and $V(\bar{x})$.

Irwin (22) and Finney (15, 16) pointed out that when the location of the dose mesh is selected at random the Spearman estimator is unbiased. This is shown as follows:

Table 4.3

Bounds for the Relative Deviation of $V(\bar{x}|x_o)$ from $\bar{V}$

Tolerance Distribution	Bound on $\frac{\lvert V(\bar{x}\lvert x_o) - \bar{V}\rvert}{\bar{V}}$			
	First Bound 1/		Second Bound 1/	
Logistic	$.4535\rho$ 2/	or $.6932\rho'$ 3/	$.2741\rho^2$ 2/	or $.6404\rho'^2$ 3/
Normal	$.4431\rho$	or $.7458\rho'$	$.2357\rho^2$	or $.6677\rho'^2$
Angular	$.4347\rho$	or $.8183\rho'$	$.1981\rho^2$	or $.7020\rho'^2$
Uniform	$.4330\rho$	or $.9000\rho'$	Not applicable	
One-Particle	$.5000\rho$	or $.6930\rho'$	Not applicable	

1/ The first bound is computed from (4.14a) and the second bound from (4.16a). (See Table 5.1 for the values of $\bar{V}$ for the several distributions.)

2/ ρ is the ratio of the dose interval to the standard deviation, $d=\rho\sigma$.

3/ ρ' is the ratio of the dose interval to distance (R) from the 20^{th} to the 80^{th} percentile, $d=\rho' R$.

$$E(\bar{x}) = \int_0^d \frac{1}{d} E(\bar{x}|x_o)\, dx_o \qquad (5.1)$$

$$= \frac{1}{d} \int_0^d \left[\sum_{-\infty}^{\infty} (x_o + d/_2 + id) \int_{x_o+id}^{x_o+id+d} dF(x) \right] dx_o$$

$$= \frac{1}{d} \sum_{-\infty}^{\infty} \int_0^d \left[(x_o + id + d/_2) \int_{x_o+id}^{x_o+(i+1)d} dF(x) \right] dx_o$$

$$= \frac{1}{d} \sum_{-\infty}^{\infty} \int_{id+d/_2}^{(i+1)d+d/_2} \left[u \int_{u-d/_2}^{u+d/_2} dF(x) \right] du$$

$$= \frac{1}{d} \int_{-\infty}^{\infty} u \int_{u-d/_2}^{u+d/_2} dF(x) \quad du$$

$$= \frac{1}{d} \int_{-\infty}^{\infty} \int_{x-d/_2}^{x+d/_2} udu \quad dF(x)$$

$$= \int_{-\infty}^{\infty} xdF(x) = \mu$$

5.3 The Mean Square Error of the Spearman Estimator

Let E_{x_o} denote the expectation with respect to x_o over interval $(0,d)$. Let $B(\bar{x}|x_o)$ denote the bias of $\bar{x}$ given x_o. The mean square error of the estimator is the variance and can be written:

$$V(\bar{x}) = E_{x_o} E\left[(\bar{x}-\mu)^2 \,|\, x_o \right]$$

$$= E_{x_o}\left[V(\bar{x}|x_o) + B^2(\bar{x}|x_o) \right]$$

$$= E_{x_o}\left[V(\bar{x}|x_o)\right] + E_{x_o}\left[B^2(\bar{x}\ x_o)\right] \qquad (5.2)$$

The first component of $V(\bar{x})$ in (5.2) can be evaluated:

$$E_{x_o}\left[V(\bar{x}|x_o)\right] = \int_o^d \frac{1}{d}\ \frac{d^2}{n} \sum_{-\infty}^{\infty} F(x_o+id)\left[1-F(x_o+id)\right]\ dx_o$$

$$= \frac{d}{n}\sum_{-\infty}^{\infty} \int_o^d F(x_o+id)\left[1-F(x_o+id)\right] dx_o$$

$$= \frac{d}{n}\sum_{-\infty}^{\infty} \int_{x_o+id}^{x_o+(i+1)d} F(x)\left[1-F(x)\right] dx$$

$$= \frac{d}{n}\int_{-\infty}^{\infty} F(x)\left[1-F(x)\right] dx \qquad (5.3)$$

Note that (5.3) is the same as (4.13), denoted by $\bar{V}$. Thus $\bar{V}$ is the average of the conditional variance of $\bar{x}$, taken over the location of the dose mesh.

A simple expression for the second component of $V(\bar{x})$ has not been obtained. The variance, $V(\bar{x})$, is written:

$$V(\bar{x}) = \frac{d}{n}\int_{-\infty}^{\infty} F(x)\left[1-F(x)\right] dx + E_{x_o}\left[B^2(\bar{x}|x_o)\right] \qquad (5.4)$$

Using the bound for the bias over values of x_o in (4.3):

$$V(\bar{x}) = \bar{V} + O(d^2)$$

For tolerance distributions satisfying the conditions given in section 4.2.7:

$$V(\bar{x}) = \bar{V} + O(d^6)$$

The second component of $V(\bar{x})$ (5.4) is independent of n and of smaller order in d than the first component, $\bar{V}$. $\bar{V}$ contains d only in the form of the factor d/n, i.e. the inverse of the number of subjects tested per unit interval on the dose scale.

5.4 Values of $\bar{V}$ for Several Tolerance Distributions

An approximation to the variance of the Spearman estimator for the case of a normal tolerance distribution, given by Gaddum (19), is equivalent to $\bar{V}$. Finney (15, 16) computed $\bar{V}$ for the normal and logistic distributions. Table 5.1 gives the values of $\bar{V}$, as proportions of σ and as proportions of the distance from the 20^{th} to 80^{th} percentile (R), for several tolerance distributions (see Table 2.2 for definitions of these tolerance distributions.)

In section 4.3 it was seen that $\bar{V}$ can be regarded as a good approximation to $V(\bar{x}|x_o)$, and that $V(\bar{x}|x_o)$ can deviate from $\bar{V}$ by at most $d^2/4n$.

In the present part $\bar{V}$ was seen to be less than the unconditional variance of $\bar{x}$, where the error is slight if d is small. In the next section $\bar{V}$ will be established as the asymptotic variance of $\bar{x}$, as defined in the same section.

6. LARGE SAMPLE PROPERTIES OF THE SPEARMAN ESTIMATOR FOR THE INFINITE EXPERIMENT

6.1. Large Sample Experiments

The experimental design for the infinite experiment consists of fixing two numbers: the number (n) of subjects

Table 5.1

Values of $\bar{V}$ for Several Tolerance Distributions

Tolerance Distribution	$\bar{V}$ 1/	
Logistic	$.5513\rho\frac{\sigma^2}{n}$ 2/	$1.2881\rho'\frac{R^2}{n}$ 3/
Normal	$.5642\rho\frac{\sigma^2}{n}$	$1.5983\rho'\frac{R^2}{n}$
Angular	$.5750\rho\frac{\sigma^2}{n}$	$2.0376\rho'\frac{R^2}{n}$
Uniform	$.5774\rho\frac{\sigma^2}{n}$	$2.4942\rho'\frac{R^2}{n}$
One-Particle	$.5000\rho\frac{\sigma^2}{n}$	$.9604\rho'\frac{R^2}{n}$

1/ $\bar{V} = \frac{d}{n}\int_{-\infty}^{\infty} F(1-F)dx$; these values were used in Table 4.3

2/ ρ is the ratio of the dose interval to the standard deviation, $d=\rho\sigma$

3/ ρ' is the ratio of the distance (R) from the 20th to the 80th percentile, $d=\rho' R$

to be tested at each of the levels, and the distance (d) between dose levels. Denote such a desgn by D(n, d). The large sample experiment is usually described in terms of a fixed d and increasing n. For the Spearman estimator this method of increasing the size of the experiment will not yield a consistent estimator. The second component of $V(\bar{x})$ in (5.4) involves the conditional bias independent of n, so that $V(\bar{x})$ does not go to zero as n goes to infinity. This points up the need for a more general concept of large sample experiments.

Let n' denote the average number of subjects tested per unit on the dose scale, $(n' = {}^{n}/_{d})$. Then the large sample experiment is obtained by letting n' go to infinity. The choice of the corresponding values for n and d will be made to minimize the mean square error for fixed n'.

6.2 Optimum Choice of n and d for the Spearman Estimator

For fixed n' and random choice of x_0 the variance of the Spearman estimator is minimized by choosing n and d as small as possible, i.e., n=1 and $d = {}^{1}/_{n'}$. This follows from the following theorem:

Theorem 6.2.1: The mean square error of the Spearman estimater based on groups of n subjects tested at dose levels d units apart is greater than the mean square error for single subjects tested at dose levels ${}^{d}/_{n}$ units apart.

Proof: Denote the two mean square errors by MSE_n and MSE_1 respectively. Denote the corresponding biases conditional on x_0 by $b_n(x_0)$ and $b_1(x_0)$ respectively.

$$MSE_n = \frac{d}{n} \int F(1-F)\,dx + E_{x_0}\,[b_n^2(x_0)] \tag{6.1}$$

$$MSE_1 = \frac{d/n}{1} \int F(1-F)\,dx + E_{x_0}\,[b_1^2(x_0)] \tag{6.2}$$

The first terms on the right hand sides of (6.1) and (6.2) are identical. Therefore it must be shown that

$$E_{x_0}[b_n^2(x_0)] > E_{x_0}[b_1^2(x_0)] \tag{6.3}$$

The conditional biases can be written:

$$b_n(x_0) = \sum_{i=-\infty}^{\infty} \int_{x_0+id}^{x_0+id+d} (x_0+id+d/2-x)\,dF \tag{6.4}$$

$$b_1(x_0) = \sum_{i=-\infty}^{\infty} \sum_{j=0}^{n-1} \int_{x_0+id+\frac{jd}{n}}^{x_0+id+\frac{(j+1)d}{n}} (x_0+id+\frac{jd}{n}+\frac{d}{2n} - x)\,dF \tag{6.5}$$

The left hand side of (6.3) is:

$$E_{x_0}[b_n^2(x_0)] = \frac{1}{d}\int_{x_0}^{x_0+d} b_n^2(x_0)\,dx_0$$

$$= \frac{1}{d} \sum_{s=0}^{n-1} \int_{x_0+\frac{sd}{n}}^{x_0+\frac{(s+1)d}{n}} b_n^2(x_0)\,dx_0$$

$$= \frac{1}{d} \sum_{s=0}^{n-1} \int_{x_0}^{x_0+\frac{d}{n}} b_n^2(x_0+\frac{sd}{n})\,dx_0 \tag{6.6}$$

The expression $b_n(x_0+\frac{sd}{n})$ appearing in the integrand of of the right hand side of (6.6) can be rewritten in terms of $b_1(x_0)$. From (6.4):

$$b_n(x_0+\frac{sd}{n}) = \sum_{i=-\infty}^{\infty} \int_{x_0+\frac{sd}{n}+id}^{x_0+\frac{sd}{n}+id+d} (x_0+\frac{sd}{n}+id+d/2-x)\,dF \quad (6.7)$$

$$S = 0,1,2,\ldots(n-1).$$

$$= \sum_{i=-\infty}^{\infty} \sum_{j=0}^{n-1} \int_{x_0+\frac{sd}{n}+id+\frac{jd}{n}}^{x_0+\frac{sd}{n}+id+\frac{(j+1)d}{n}} (x_0+\frac{sd}{n}+id+d/2-x)\,dF \quad (6.8)$$

Then, adding and subtracting $\frac{jd}{n}+\frac{d}{2n}$ in the integrand of (6.8) :

$$b_n(x_0+\frac{sd}{n}) = \sum_{i=-\infty}^{\infty} \sum_{j=0}^{n-1} \int_{x_0+id+\frac{(j+s)d}{n}}^{x_0+id+\frac{(j+s+1)d}{n}} [x_0+id+\frac{(j+s)d}{n}+\frac{d}{2n}-x]\,dF$$

$$+\frac{d}{2} \sum_{i=-\infty}^{\infty} \sum_{j=0}^{n-1} \int_{x_0+id+\frac{(j+s)d}{n}}^{x_0+id+\frac{(j+s+1)d}{n}} (1-\frac{2j+1}{n})\,dF \quad (6.9)$$

From (6.5), expression (6.9) can be written:

$$b_n(x_0+\frac{sd}{n}) = b_1(x_0+\frac{sd}{n}) + \frac{d}{2}\sum_{j=0}^{n-1}(1-\frac{2j+1}{n})A_{j+s} \tag{6.10}$$

$$\text{where } A_{j+s} = \sum_{i=-\infty}^{\infty}\int_{x_0+id+\frac{(j+s)d}{n}}^{x_0+id+\frac{(j+s+1)d}{n}} dF$$

$$\text{Note that } \sum_{s=0}^{n-1} A_{j+s} = 1 \quad \text{for any } j \tag{6.11}$$

Since b_1 is periodic with period $^d/_n$ it follows from (6.10) that

$$b_n(x_0+\frac{sd}{n}) = b_1(x_0) + \frac{d}{2}\sum_{j=0}^{n-1}(1-\frac{2j+1}{n})A_{s+j} \tag{6.12}$$

Substituting (6.12) in the right hand side of (6.6):

$$E_{x_0}[b_n^2(x_0)] = \frac{1}{d}\sum_{s=0}^{n-1}\int_{x_0}^{x_0+\frac{d}{n}}[b_1^2(x_0) + \frac{d}{2}\sum_{j=0}^{n-1}(1-\frac{2j+1}{n})A_{s+j}]^2\, dx_0 \tag{6.13}$$

On squaring the expression in the integrand in (6.13) the middle term will be :

$$\sum_{s=0}^{n-1} \int_{x_0}^{x_0+\frac{d}{n}} \sum_{j=0}^{n-1} \left(1- \frac{2j+1}{n}\right) A_{s+j} \, dx_0 \tag{6.14}$$

Summing first on s the sum of the A_{j+s} is one and the sum with respect to j will then be zero. Therefore (6.13) becomes:

$$E_{x_0}[b_n^2(x_0)] = E_{x_0}[b_1^2(x_0)] + \\ + \frac{d}{4} \int_{x_0}^{d/n} \sum_{s=0}^{n-1} \left(\sum_{j=0}^{n-1} \left(1 - \frac{2j+1}{n}\right) A_{j+s}\right)^2 dx_0 \tag{6.15}$$

The second term on the right hand side of (6.15) cannot be zero if F is a distribution function. Therefore,

$$E_{x_0}[b_n^2(x_0)] > E_{x_0}[b_1^2(x_0)] \qquad \text{Q.E.D.}$$

6.3 Large Sample Properties of the Spearman Estimator

If n' is increased, with n=1 and $d={}^1/_{n'}$, as required, it follows from (5.4) and the fact that the bias squared has a bound of order d^2, that the estimator is consistent and the variance is

$$V(\bar{x}) = \frac{1}{n'} \int_{-\infty}^{\infty} F(1-F)\,dx + O({}^1/_{n'^2}) \qquad (6.9)$$

It is convenient, in the case of a sequence of random variables, to approximate the variances by simpler terms correct to order n^{-1}. (When the sequence does not have variances, the variances of a sequence of limiting distributions may be used.) Such a sequence of approximations will be called the asymptotic variances. In this sense the first term on the right of (6.9) will be called the asymptotic variance of $\bar{x}$ and will be denoted by $V_A(\bar{x})$.

$$V_A(\bar{x}) = \frac{1}{n'} \int_{-\infty}^{\infty} F(1-F)\,dx \qquad (6.10)$$

7. LARGE SAMPLE EFFICIENCY OF THE SPEARMAN ESTIMATOR FOR THE INFINITE EXPERIMENT

7.1 Previous Comparisons of the Spearman Estimator with the Maximum Likelihood Estimator

Finney (15,16) computed the asymptotic variances of the maximum likelihood estimator, averaged over choices of x_o, for the finite experiment, for the normal and logistic tolerance distributions, and then took the limit as the number of levels went to infinity. He compared these values with the mean square error of the Spearman estimator over choices of x_o for the same two distributions. The ratios were .9814 and 1.0000 respectively.

Cornfield and Mantel (10) showed that for the logistic tolerance distribution, the maximum likelihood estimator and the Spearman estimator were approximately equal and this algebraic approximation improved as $d \rightarrow 0$. Bross (9) evaluated some sampling distributions through enumeration for the maximum likelihood estimator and the Spearman estimator. He used the logistic tolerance distribution, four dose levels, with $n=2$ and also $n=5$. In all cases examined, the Spearman estimator was concentrated more closely about the true mean tolerance than was the maximum likelihood estimator. These computational results were reproduced by Haley (21) for the normal tolerance distribution.

These results indicate that the Spearman estimator compares favorably in precision with the maximum likelihood estimator, at least for the normal and logistic distributions. In this section the asymptotic efficiency of the Spearman estimator is defined and various results are reported concerning tolerance distributions that minimize or maximize this efficiency. Efficiencies for the common tolerance distributions are given, the values for the logistic and normal being the same as the ratios given by Finney.

7.2 Definition of Asymptotic Efficiency

The efficiency of an estimator can be defined in terms of the quantity, I, called the information:

$$I = E\left(\frac{\partial \ln f}{\partial \mu}\right)^2$$

where f is the frequency function for the random variables on which the estimator depends (11).

For the infinite experiment with random choice of x_o, the information is (see Appendix I):

$$I = \frac{n}{d} \int_{-\infty}^{\infty} \frac{F_{\mu}^{2}(x)}{F(x)\left[1-F(x)\right]} dx$$

The asymptotic efficiency, E, of an estimator for the infinite experiment will be defined as the ratio of $^1/_I$ to the asymptotic variance of the estimator. For the Spearman estimator,

$$E = \frac{1/I}{V_A(\bar{x})}$$

$$= \left[\int_{-\infty}^{\infty} F(x)\left[1-F(x)\right] dx \int_{-\infty}^{\infty} \frac{F_\mu^{\,2}(x)}{F(x)\left[1-F(x)\right]}\, dx \right]^{-1} .$$

$1/I$ is the asymptotic variance $(n \to \infty)$ of the maximum likelihood estimator, so that E measures the efficiency of the Spearman estimator relative to the maximum likelihood estimator.

7.3 The Spearman Efficiency for Several Tolerance Distributions

The following sections present computational results for specific tolerance distributions. The results are summarized in Table 7.1.

7.3.1. Logistic

The efficiency for the logistic is 1.0 since the logistic distribution satisfies

$$F_x^{\,2} \equiv F_\mu^{\,2}$$

$$\beta\ F(1-F) = F_x$$

7.3.2. Normal

Finney evaluated the efficiency for the normal tolerance distribution. The integral of $F_\mu^{\,2}/F(1-F)$ has to be obtained through numerical methods. The result is E=.9814

7.3.3. Angular

For the angular distribution, both integrals are easily evaluated and E=.8106.

7.3.4. Uniform

The definition of E is not applicable to the uniform distribution since the regularity conditions used in obtaining I are not fulfilled. The particular difficulty is that the distribution function is not differentiable for all values of μ (see Appendix I).

7.3.5. One Particle

The integrals for the one-particle distribution are:

$$\frac{d}{n}\int_0^{\infty} F(1-F)dx = \frac{d}{n}\frac{1}{2}\sigma$$

$$\frac{n}{d}\int_0^{\infty} \frac{F_{\mu}^{\;2}}{F(1-F)}\,dx = \frac{n}{d}\,\sigma\int_0^{\infty}\frac{y^2e^{-y}dy}{1-e^{-y}} = \frac{n}{d}\cdot\sigma\sum_1^{\infty}\frac{1}{k^3} = \frac{2.4043n\sigma}{d}$$

and the efficiency is .83191...

7.3.6. Algebraic

The above examples all involve distributions with first moments for all values of the parameter. The algebraic distribution (Table 2.2) does not have a first moment for all values of the one parameter:

$$F(x;\, s) = 1-x^{-s} \qquad s > 1,\ x \geq 1$$

This distribution has a first moment with value $\mu = \frac{s}{s-1}$ if $s > 1$. For $s \leq 1$ the first moment does not exist.

$$\frac{d}{n}\int_1^{\infty} F(1-F)dx = \frac{d}{n}\,\frac{s}{(s-1)(2s-1)} \qquad s > 1$$

$$\frac{n}{d}\int_{-\infty}^{\infty} \frac{F_\mu^{\,2}}{F(1-F)}\,dx = \frac{n}{d}\,2(s-1)^4 \sum_{k=1}^{\infty} \frac{1}{(ks-1)^3} \qquad s > 1$$

and

$$E = \frac{2-{}^1/_s}{2} \quad \frac{1}{\sum\limits_{k=1}^{\infty}\left[\frac{1-{}^1/_s}{k-{}^1/_s}\right]^3} .$$

Thus, in contrast with the preceding examples, E depends on s.

$$\lim_{s\to 1} E = \frac{1}{2}\lim_{s\to 1}\,(2-{}^1/s)\;\frac{1}{\lim\limits_{s\to 1}\sum\limits_{k=1}^{\infty}\frac{(1-{}^1/s)^3}{(k-{}^1/s)^3}}$$

The limit in the denominator is one since:

$$\lim_{s\to 1}\sum_{k=1}^{\infty}\frac{(s-1)^3}{(ks-1)^3} = \lim_{s\to 1}\left(1+\sum_{k=2}^{\infty}\frac{(s-1)^3}{(ks-1)^3}\right) = 1$$

Also

$$\lim_{s\to\infty} E = \frac{1}{\sum\limits_{1}^{\infty}\frac{1}{k^3}} = .8319\ldots$$

Note that the limit of the efficiency as s becomes large is identical with the efficiency for the one-particle tolerance distribution. This might be anticipated by rewriting the two distributions from Table 2.2.

Algebraic Distribution: $1-(x+1)^{-s}$ $\quad x \geq 0 \quad \mu = \frac{1}{s-1}$

One-Particle Distribution: $1-e^{-(s-1)x}$ $\quad x \geq 0 \quad \mu = \frac{1}{s-1}$

The ratio of the i^{th} moments of the two distributions is

$$\frac{(s-1)(s-2)\ldots(s-i)}{(s-1)^i} .$$

This ratio goes to one as s becomes large.

7.4 The Logistic Tolerance Distribution and the Spearman Estimator

Theorem 7.4.1:

The logistic tolerance distribution is the only symmetrical tolerance distribution, with a translation parameter as the single unknown parameter, for which the Spearman estimator has full efficiency.

Proof: Let μ be the translation parameter and let the tolerance distributions be written $F(y-\mu)$. The efficiency for a given distribution, F, is:

$$E(F)=\left[\int_{-\infty}^{\infty}\frac{f^2(x)}{F(x)\left[1-F(x)\right]}dx\int_{-\infty}^{\infty}F(x)\left[1-F(x)\right]dx\right]^{-1}$$

Let G be an extremal function (symmetrical, differentiable) of the functional $Y(F)=\left[E(F)\right]^{-1}$. Let $V(x)$ be any function satisfying the conditions:

$V(x)=V(-x)$ (7.1)

$V(x)$ is differentiable for all x (7.2)

for all t in a neighborhood of $t=0$, $G(x)+tV(x)$ is a distribution function with first moment. (7.3)

Then $y(t)=Y\left[G(x)+tV(x)\right]$ is a function of t differentiable at $t=0$, and $y'(0)=0$.

Table 7.1

Large Sample Efficiency (E) of the Spearman Estimator for the Infinite Experiment for Several Tolerance Distributions 1/

TOLERANCE DISTRIBUTION	EFFICIENCY 2/
1. Logistic	1.0000
2. Normal	.9814
3. Angular	.8106
4. One-Particle	.8319
5. Algebraic	$.500 \leq E \leq .8319$

1/ $$E = \left[\int F(1-F)\,dx \int \frac{F_\mu^2}{F(1-F)}\,dx \right]^{-1}$$

2/ See sections 7.3.1, 7.3.2, 7.3.3, 7.3.5 and 7.3.6 for computations.

Let $G'(x)$ be denoted by $g(x)$ and $V'(x)$ be denoted by $v(x)$. Then

$$y'(o) = \int_{-\infty}^{\infty} \frac{g^2}{G(1-G)}\, dx \int_{-\infty}^{\infty} V(1-2G)\, dx$$
$$+ \int_{-\infty}^{\infty} \frac{2G(1-G)gv - Vg^2(1-2G)}{G^2(1-G)^2}\, dx \int_{-\infty}^{\infty} G(1-G)\, dx$$

Since G has a symmetrical density and V is symmetrical, the integrand of the following integral is an odd function and

$$\int_{-\infty}^{\infty} \frac{2G(1-G)gv}{G^2(1-G)^2}\, dx = 0$$

Then

$$y'(o) = \int_{-\infty}^{\infty} V(1-2G) \left[C_1 - \frac{C_2 g^2}{G^2(1-G)^2} \right] dx = 0$$

where

$$C_1 = \int_{-\infty}^{\infty} \frac{g^2}{G(1-G)}\, dx$$

$$C_2 = \int_{-\infty}^{\infty} G(1-G) dx$$

$V(x)$ can be any function satisfying conditions (7.1), (7.2) and (7.3) and $1-2G(x)$ cannot be identically zero on the infinite interval. Therefore the necessary condition for $G(x)$ to be an extremal function of $Y(F)$ is that

$$C_1 \frac{\sigma_2 g^2(x)}{\left\{ G(x) \left[1-G(x)\right]\right\}^2} \equiv 0$$

This implies that G(x) is of the logistic form:

$$G(x) = \left[1+e^{-(\alpha+\beta x)}\right]^{-1}$$

Q.E.D.

7.5 Distributions with Efficiency of the Spearman Estimator Close to Zero

If F(x) has a first moment the variance of the Spearman estimator exists, i.e. the integral with respect to F(1-F) is finite. Then, if the information is finite, the efficiency of the Spearman estimator is greater than zero.

Distributions with a translation parameter as the single unknown parameter can be specified for which the Spearman efficiency is arbitrarily close to zero. Consider,

$$F_e(x;\mu) = K(e)\int_{-\infty}^{x} \frac{dt}{\left[1+\frac{(t-\mu)^2}{1+2e}\right]^{1+e}} \qquad \begin{array}{l} -\infty < x < \infty \\ -\infty < \mu < \infty \\ 0 < e \end{array}$$

The efficiency can be made arbitrarily small by choosing e close to zero; V(e) is unbounded as e goes to zero, while I(e) is bounded away from zero.

Consider the following bound for V(e):

$$V(e) = \frac{d}{n}\int_{-\infty}^{\infty} F_e(x;\mu)\left[1-F_e(x;\mu)\right] dx$$

$$= \frac{d}{n} 2\int_{0}^{\infty} F_e(x;o) \left[1-F_e(x;o)\right] dx$$

$$\geqq \frac{d}{n} \int_{0}^{\infty} \left[1-F_e(x;o)\right] dx$$

$$\geqq \frac{d}{n} \int_{0}^{\infty} \int_{x}^{\infty} \frac{K(e)dt}{\left[1+ \frac{t^2}{1+2e}\right]^{1+e}} dx$$

$$\geqq \frac{d}{n} \int_{1}^{\infty} \int_{x}^{\infty} \frac{K(e)dt}{\left[1+t^2\right]^{1+e}} dx$$

$$\geqq \frac{dK(e)}{n2^{1+e}} \int_{1}^{\infty} \int_{x}^{\infty} \frac{dt}{t^{2+2e}} dx$$

$$\geqq \frac{dK(e)}{n2^{1+e}} \frac{1}{1+2e} \frac{1}{2e} \tag{7.4}$$

The constant $K(e)$ necessary to make F_e a distribution is greater than $\frac{1}{3}$ for $e < 1$; therefore $V(e)$ goes to infinity as e goes to zero.

$I(e)$ can be written:

$$I(e) = \int_{-\infty}^{\infty} \frac{\left[\frac{\partial F_e}{\partial} \frac{(x;\mu)}{\mu}\right]^2}{F_e(x;\mu) \left[1-F_e(x;\mu)\right]} dx$$

To show that $I(e)$ is bounded away from zero as e goes to zero, first note that

$$\left[\frac{\partial F}{\partial \mu}\right]^2 = \left[\frac{\partial F}{\partial x}\right]^2 = \frac{K^2(e)}{\left[1+\frac{(x-\mu)^2}{1+2e}\right]^{2+2e}}$$

Using the symmetry of the integrand and $F(1-F) \leq \frac{1}{4}$:

$$I(e) = 2\int_{\mu}^{\infty} \frac{\left[\frac{\partial F}{\partial \mu}\right]^2}{F(1-F)}\, dx$$

$$\geq 8\int_{0}^{\infty} \frac{K^2(e)dx}{\left[1+\frac{x^2}{1+2e}\right]^{2+2e}}$$

Then the following inequalities are obtained:

$$I(e) \geq 8K^2(e)\int_{1}^{\infty} \frac{dx}{\left[1+x^2\right]^{2+2e}}$$

$$\geq 8K^2(e)\ \frac{1}{2^{2+2e}}\int_{1}^{\infty} \frac{dx}{x^{4+4e}}$$

$$\geq 8K^2(e)\ \frac{1}{2^{2+2e}}\ \frac{1}{3+4e}$$

$$\geq \frac{8}{3^2}\ \frac{1}{2^{2+2e}}\ \frac{1}{3+4e} \qquad (7.5)$$

It follows from (7.4) and (7.5) that $E(e)$ goes to zero as e goes to zero.

7.7 Two Parameter Families of Tolerance Distributions

The results of the previous paragraphs are applicable without modification to the case of scale parameter unknown when the tolerance distribution is symmetrical. (Estimation of the scale parameter itself is duscussed in Appendix III.)

Let the tolerance distribution be of the form $F(y)$ where $y = \beta(x-\mu)$. Let both β and μ be unknown. The infinite experiment information matrix is given in Appendix I.

Letting

$$A_1 = \frac{\int_{-\infty}^{\infty} tWdt}{\int_{-\infty}^{\infty} Wdt}$$

$$A_2 = \frac{\int_{-\infty}^{\infty} (t-A_1)^2 Wdt}{\int_{-\infty}^{\infty} Wdt},$$

where $W(t) = \frac{\left[F'(t)\right]^2}{F(t)\left[1-F(t)\right]}$.

The inverse element, I^{11}, corresponding to μ is

$$I^{11} = \frac{d}{n\beta \int_{-\infty}^{\infty} W(t)dt} \; \frac{A_1^2 + A_2}{A_2}$$

Note that if $A_1=0$, then $I^{11} = {}^1/_I$ when I is the information for scale parameter known. If F is a symmetrical distribution then $W(t)$ is symmetrical and $A_1=0$.

Note also that if $A_1 \neq 0$, the effect is to increase the value of I^{11} above the corresponding value for scale parameter known. The variance of the Spearman estimation is unchanged. Hence, the efficiency for the Spearman estimator would be greater in such cases for scale parameter unknown than for scale parameter known.

8. THE SPEARMAN ESTIMATOR FOR THE FINITE EXPERIMENT

8.1 Finite Experiments

In previous sections (3, 4, 5, 6, 7) the range of experimentation was infinite. The results obtained are useful in designing and interpreting experiments in which the dose levels cover the greater portion of the range of F(x), say from .01 to .99. As a supplement to these results it is of interest to investigate the effect of using a finite set of dose levels.

Let x_o be an a priori estimate of μ and let the experiment involve 2k+1 dose levels regularly spaced over the interval $(x_o-a,\ x_o+a)$. The dose levels are $x_i=x_o+id$, $i=0, \pm1, \pm2, \ldots \pm k$, with $kd=a$. Let N be the total number of subjects used in the assay, $N=(2k+1)n$.

8.2 The Spearman Estimator

The variance and bias of the Spearman estimator are

$$V_a(\bar{x}|x_o) = \frac{d^2}{n} \sum_{-k}^{k} F_i(1-F_i) \qquad (8.1)$$

$$B_a(\bar{x}|x_o) = (x_o-a-{}^d/_2)F(x_o-a-{}^d/_2)+ \sum_{-k}^{k-1} (x_i+{}^d/_2)(F_{i+1}-F_i)$$

$$+(x_o+a+{}^d/_2)\left[1-F(x_o+a+\frac{d}{2})\right] -\mu \qquad (8.2)$$

<u>Theorem 8.2.1</u> Let the range of dose levels be (x_o-a, x_o+a), the total sample size be N, and the numbers of subjects at each dose level be equal (n). Then the maximum variance of $\bar{x}$ over all possible F is minimized by minimizing d (i.e. by maximizing the number of dose levels).

<u>Proof:</u> From (8.1) the variance of the estimator is

$$V_a(\bar{x} \mid x_o) = \left(\frac{a}{k}\right)^2 \frac{(2k+1)}{N} \sum_{-k}^{k} F_i(1-F_i) \qquad (8.3)$$

$F_i(1-F_i) \leq \frac{1}{4}$ so that

$$V_a(\bar{x} \mid x_o) \leq a^2 \frac{(2k+1)^2}{k^2} \frac{1}{4N}$$

The bound is attained for $F(x)$ a two point distribution defined by:

$$P(x=x_o-a-e) = {}^1/_2 \qquad e > o$$

$$P(x=x_o+a+e) = {}^1/_2$$

The bound is minimized by choosing k as large as possible, i.e. by choosing d as small as possible. Q.E.D.

There are distributions for which an increase in k results in an <u>increase</u> in the variance of $\bar{x}$. Consider the distribution given by:

$$P(x=x_o-{}^3/_4 a)={}^1/_2$$

$$P(x=x_o+{}^3/_4 a)={}^1/_2$$

The variances for three and five levels are:

$$k=1 : V_a(\bar{x} \mid x_o) = \frac{3a^2}{4N}$$

$$k=2 : V_a(\bar{x} \mid x_o) = \frac{15a^2}{16N} .$$

The variance for the normal distribution for several values of k is shown in Table 8.1.a. The results are for $x_o=\mu$. For the dose ranges used, the variance decreases as k increases.

For asymmetrical location ($x_o \neq \mu$) the Spearman estimator will be biased. Table 8.1.b presents the mean square error for the normal distribution for several values of a, k, and N, for several values of x_o.

8.3 Information for the Finite Experiment, Scale Parameter Known

Denote the information for the finite experiment described in section 8.1 by $I_a(x_o)$. Then

$$I_a(x_o) = \frac{N}{2k+1} \sum_{-k}^{k} \frac{F_\mu^2 \left(\frac{ia}{k}\right)}{F\left(\frac{ia}{k}\right) \left[1-F\left(\frac{ia}{k}\right)\right]}$$

Table 8.2.a presents the values of $^N/_{\sigma^2} I_a(x_o)$ for the normal distribution, for $x_o=\mu$, for several values of a and k. The results show that there are extreme situations (three levels placed at $\mu-10\sigma$, μ, and $\mu+10\sigma$) for which an increase in k results in a decrease in the information. However, when a is 2σ or less, the denser the dose levels,

Table 8.1.a

Variance (V_a) of the Spearman Estimator for the Mean of a Normal Tolerance Distribution, for the Finite Experiment

Middle Dose Location $(\frac{x_o-\mu}{\sigma})$	Dose Range $(\pm \frac{a}{\sigma})$	$\frac{NV_a(\bar{x} \mid x_o)}{\sigma^2}$						
		k=1[1]	k=2	k=3	k=4	k=8	k=20	k=40
0	1/2	.51	.36	.32	.30	.27	.25	.24
0	1	1.55	1.18	1.06	1.00	.91	.87	.83
0	2	3.54	2.81	2.58	2.51	2.36	2.25	2.19
0	4	12.00	5.89	5.08	5.02	4.75	4.58	4.50
0	10	75.00	31.25	19.44	15.18	12.37	11.46	11.23

1/ The number of dose levels is 2k+1.

Table 8.1.b

Mean Square Error (MSE_a) of the Spearman Estimator for the Mean of the Normal Tolerance Distribution, for the Finite Experiment

Middle Dose Location $(\frac{x_o-\mu}{\sigma})$	Dose Range $(\pm \frac{a}{\sigma})$	$N\left[\frac{MSE_a(\bar{x} \mid x_o)}{\sigma^2}\right]$ N=10 k=1[1]	N=10 k=2	N=10 k=4	N=100 k=1[1]	N=100 k=2	N=100 k=4
0	±1	1.55	1.18		1.55	1.18	
.5		1.51	1.22		1.92	2.36	
1.0		1.55	1.63		4.51	8.60	
2.0		5.13	7.99		47.09	77.35	
3.0		23.30	31.20		232.36	311.65	
4.0		62.57	75.67		625.65	756.66	
0	±2	3.54	2.81		3.54	2.81	
.5		3.38	2.79		3.38	2.79	
1.0		3.22	2.71		3.22	2.76	
2.0		3.29	2.37		3.48	5.38	
3.0		2.64	5.45		11.86	47.41	
4.0		11.20	23.35		109.60	232.41	
5.0		40.12	62.57		401.06	625.66	
0	±4		5.89	5.08		5.89	5.08
.5			5.64	5.08		5.64	5.08
1.0			5.39	5.08		5.39	5.08
2.0			5.89	5.07		5.89	5.07
3.0			5.37	4.87		5.37	4.93
4.0			5.47	4.00		5.66	7.00
5.0			3.72	6.08		12.94	48.03
6.0			11.38	23.44		109.78	232.50
7.0			40.13	62.58		401.07	625.66

[1] The number of dose levels is 2k+1.

Table 8.2.a

Information (I_a) for Estimation of the Mean of the Normal Tolerance Distribution with Scale Parameter Known, for the Finite Experiment

Middle Dose Location $(\frac{x_0-\mu}{\sigma})$	Dose Range $(\pm \frac{a}{\sigma})$	$\frac{N}{\sigma^2 I_a(x_0)}$ k=1[1/]	k=2	k=3	k=4	k=8	k=20	k=40
0	1/2	1.67	1.64	1.63	1.63	1.62	1.62	1.62
0	1	1.98	1.87	1.83	1.82	1.79	1.78	1.77
0	2	3.34	2.82	2.66	2.59	2.48	2.42	2.38
0	4	4.70	5.56	5.16	4.98	4.71	4.54	4.44
0	10	4.71	7.85	10.81	12.22	11.76	11.35	11.08

[1/] The number of dose levels is 2k+1.

i.e. the greater the number of dose levels, the greater the information.

Table 8.2.b presents the values of $N/\sigma^2 I_a(x_o)$ for values of a, k and N, for values of $x_o \neq \mu$.

8.4 Efficiency of the Spearman Estimator for the Finite Experiment, Scale Parameter Known

The efficiency [1] of the Spearman estimator for the finite experiment will be defined as the ratio of the inverse of the information to the mean square error of the estimator:

$$E_a(x_o) = \frac{1/I_a\ (x_o)}{MSE_a(\bar{x} \mid x_o)} \qquad (8.4)$$

Table 8.3.a presents computational results for the normal distribution for $x_o = \mu$ for several values of a and k.

Table 8.3.b presents computational results for the normal distribution for several values of a, k and N, for values of $x_o \neq \mu$.

8.5 Efficiency of the Spearman Estimator for the Finite Experiment, Scale Parameter Unknown

Table 8.3.b indicates that the efficiency of the Spearman

---

[1] The efficiency (E_a) as defined in terms of information (8.4) is a useful measure because the information is intrinsic to the experiment itself and not dependent on any method of estimation. The inverse of the information cannot be taken as an absolute lower bound on the variances of all estimators, nor can it be assumed that there is any estimator with variance this small. However, it is a lower bound for the variances of all unbiased estimators.

Table 8.2.b

Information (I_a) for Estimation of the Mean of the Normal Tolerance Distribution with Scale Parameter Known, for the Finite Experiment

Middle Dose Location $(\frac{x_o-\mu}{\sigma})$	Dose Range $(\pm \frac{a}{\sigma})$	$\frac{N}{\sigma^2 I_a(x_o)}$ k=1[1]	k=2	k=4
0	±1	1.98	1.87	
.5		2.10	2.00	
1.0		2.49	2.43	
2.0		5.13	5.53	
3.0		20.51	25.08	
4.0		197.37	270.27	
0	±2	3.34	2.82	
.5		3.33	2.86	
1.0		3.36	3.01	
2.0		3.90	4.09	
3.0		6.62	8.55	
4.0		22.78	34.18	
5.0		205.48	328.95	
0	±4		5.56	4.98
.5			5.54	4.98
1.0			5.52	4.98
2.0			5.56	5.02
3.0			5.61	5.42
4.0			6.51	7.37
5.0			11.03	15.39
6.0			37.97	61.52
7.0			342.47	592.11

[1] The number of dose levels is 2k+1.

estimator can be very small for $N=100$ when the a priori estimate, x_o, for μ is in error. The information was computed assuming scale parameter known. If the scale parameter is unknown the information is considerably decreased for $x_o \neq \mu$. Table 8.4 presents the element $I_a^{11}(x_o)$ of the inverse of the information matrix for estimation of the mean of the normal tolerance distribution for both location and scale parameter unknown. Table 8.5 presents the corresponding efficiency (E_a^{11}) of the Spearman estimator for two unknown parameters. Tables 8.3.a, 8.3.b and 8.5 demonstrate that for the usual finite level design with limited numbers of subjects the Spearman estimator has high efficiency relative to the information in the experiment, when the tolerance distribution is normal.

9. REGULAR BEST ASYMPTOTICALLY NORMAL ESTIMATORS WITH THE WRONG MODEL

9.1 General Discussion

One advantage of the Spearman estimator is that no parametric form need be specified for the tolerance distribution. This advantage would be of no practical importance if the competing parametric estimator based on a common model has a distribution that is insensitive to moderate changes in the functional form of the true tolerance distribution. In this section computations are presented to illustrate the effect

Table 8.3.a

Efficiency of the Spearman Estimator for the Mean of the Normal Tolerance Distribution with Scale Parameter known, for the Finite Experiment

Middle Dose Location	Dose Range	$E_a(x_o)$[1]						
$(\frac{x_o-\mu}{\sigma})$	$(\pm \frac{a}{\sigma})$	k=1[2]	k=2	k=3	k=4	k=8	k=20	k=40
0	1/2	3.27	4.56	5.09	5.43	6.00	6.48	6.75
0	1	1.28	1.58	1.73	1.82	1.97	2.04	2.13
0	2	.94	1.00	1.03	1.03	1.05	1.08	1.09
0	4	.39	.94	1.02	.99	.99	.99	.99
0	10	.06	.25	.56	.80	.95	.99	.99

[1] $E_a(x_o) = \dfrac{1}{I_a(x_o)MSE_a(\bar{x} \mid x_o)}$

[2] The number of dose levels is 2k+1.

Table 8.3.b

Efficiency (E_a) of the Spearman Estimator for the Mean of the Normal Tolerance Distribution with Scale Parameter Known, for the Finite Experiment

Middle Dose Location $(\frac{x_o-\mu}{\sigma})$	Dose Range $(\pm \frac{a}{\sigma})$	$E_a(x_o)$ 1/ N=10 k=1 2/	N=10 k=2	N=10 k=4	N=100 k=1	N=100 k=2	N=100 k=4
0	±1	1.28	1.58		1.28	1.58	
.5		1.39	1.64		1.09	.85	
1.0		1.61	1.49		.55	.28	
2.0		1.00	.69		.11	.07	
3.0		.88	.80		.09	.08	
4.0		3.15	3.57		.32	.36	
0	±2	.94	1.00		.94	1.00	
.5		.99	1.02		.99	1.02	
1.0		1.04	1.11		1.04	1.09	
2.0		1.19	1.73		1.12	.76	
3.0		2.51	1.57		.56	.18	
4.0		2.03	1.46		.21	.15	
5.0		5.12	5.26		.51	.53	
0	±4		.94	.98		.94	.98
.5			.98	.98		.98	.98
1.0			1.02	.98		1.02	.98
2.0			.94	.99		.94	.99
3.0			1.04	1.11		1.04	1.10
4.0			1.19	1.84		1.15	1.05
5.0			2.97	2.53		.85	.32
6.0			3.34	2.62		.36	.26
7.0			8.53	9.46		.85	.95

1/ $E_a(x_o) = \frac{1}{I_a(x_o) MSE_a(\bar{x} \mid x)}$

2/ The number of dose levels is 2k+1.

Table 8.4

Inverse Information (I_a^{11}) for Estimation of the Mean of the Normal Tolerance Distribution with Scale Parameter Unknown, for the Finite Experiment

Middle Dose Location $(\frac{x_o-\mu}{\sigma})$	Dose Range $(\pm \frac{a}{\sigma})$	$\frac{NI_a^{11}(x_o)}{\sigma^2}$ 1/ k=1 2/	k=2	k=4
0	±1	1.98	1.87	
.5		2.40	2.53	
1.0		4.31	5.43	
2.0		38.63	51.20	
3.0		907.36	959.53	
0	±2	3.34	2.82	
.5		3.33	2.87	
1.0		3.37	3.13	
2.0		4.71	6.98	
3.0		66.80	62.94	
0	±4		5.56	4.98
.5			5.54	4.98
1.0			5.52	4.98
2.0			5.56	5.03
3.0			5.62	5.63
4.0			7.85	12.57
5.0			111.33	113.30

1/ $I_a^{11}(x_o)$ is the element of the inverse of the information matrix corresponding to the estimator of μ .

2/ The number of dose levels is (2k+1).

Table 8.5

Efficiency (E_a^{11}) of the Spearman Estimator for the Mean of the Normal Tolerance Distribution with Scale Parameter Unknown, for the Finite Experiment

Middle Dose Location $(\frac{x_o-\mu}{\sigma})$	Dose Range $(\pm \frac{a}{\sigma})$	$E_a^{11}(x_o)$ [1]					
		N=10			N=100		
		k=1 [2]	k=2	k=4	k=1	k=2	k=4
0	±1	1.28	1.58		1.28	1.58	
.5		1.59	2.08		1.25	.88	
1.0		2.79	3.34		.96	.39	
2.0		7.52	6.41		.82	.66	
3.0		38.94	30.76		3.90	3.08	
0	±2	.94	1.00		.94	1.00	
.5		.99	1.03		.99	1.03	
1.0		1.05	1.15		1.05	1.12	
2.0		1.43	2.95		1.35	2.53	
3.0		25.29	26.56		5.63	11.55	
0	±4		.94	.98		.94	.98
.5			.98	.98		.98	.98
1.0			1.02	.98		1.02	.98
2.0			.94	.99		.94	.99
3.0			1.05	1.16		1.05	1.14
4.0			1.44	3.14		1.39	1.79
5.0			29.93	18.65		8.61	2.36

[1] $E_a^{11}(x_o) = \dfrac{I_a^{11}(x_o)}{MSE_a(\tilde{x} \mid x_o)}$

[2] The number of dose levels is 2k+1.

on some RBAN estimators[1] due to changes in the functional form of the tolerance distribution.

The angular model is used and the characteristics of some RBAN estimators based on this model are examined for true tolerance distributions with the forms: logistic, normal and uniform. The experimental designs for which computations are given are one, two, and five level designs with scale parameter known, and some two level designs with scale parameter unknown.

The RBAN estimator used in each of these finite designs is an explicit function of the independent binomial variates corresponding to the several dose levels. Consequently the mean and variance of the limiting normal distribution of the estimator (as the sample sizes at the fixed dose levels increase) can be computed. These values are called the asymptotic mean and asymptotic variance. Since the estimator is inconsistent when the wrong model is used, the asymptotic mean square error is computed from the asymptotic mean and variance and this value is compared with the asymptotic variance of the RBAN estimator under the correct model.

9.2 One Level Experiment

The model used for the tolerance distribution is the angular distribution (see Table 2.2). Assume that the scale

---

[1] See Neyman (26) for the definition of RBAN (regular best asymptotically normal) estimators. See Taylor (29) for a discussion of RBAN estimators in bioassay.

parameter, β, is known and that the experiment consists of testing N subjects at the dose level, $x=0$. Let the observed proportion responding at $x=0$ be denoted by p and the expected proportion be denoted by P. Let $y=\sin^{-1}\sqrt{p} - \pi/4$. Then the maximum likelihood estimator of μ is:

$$\mu_1^* = -\frac{y}{\beta} \qquad (9.1)$$

The asymptotic mean, variance and mean square error of the estimator, for a given value of P, are:

$$E_{aA}(\mu_1^*) = -\frac{Y}{\beta} \qquad (9.2)$$

$$V_{aA}(\mu_1^*) = \frac{1}{\beta^2}\left(\frac{1}{4N}\right) \qquad (9.3)$$

$$MSE_{aA}(\mu_1^*) = \frac{1}{\beta^2}\left(\frac{1}{4N}\right)+\left(\mu+\frac{Y}{\beta}\right)^2 \qquad (9.4)$$

where $Y=\sin^{-1}\sqrt{P} - \pi/4$

and $P=E(p)$.

If the true tolerance distribution is angular with scale parameter β, then $\mu = -\frac{Y}{\beta}$ and the asymptotic mean square error is the variance (9.3). Denote the angular distribution by G. If the true tolerance distribution is $F\neq G$, then the estimator will not be consistent. The asymptotic variance will remain the same but the bias contribution to the mean square error will not be zero.

The asymptotic mean square error can be computed for any given F and N. Denote this value by $MSE_{aA}(\mu^* \mid F)$. In specifying F it is necessary to choose the value of the scale parameter. This should be done so that the F is "comparable" to the model, G, with its known, fixed β . Three methods for choosing scale parameters for the tolerance distributions are used: (i) equating standard deviations. (ii) equating the distances between two specified percentiles. (iii) equating the information per observation.

As a measure of the effect of the tolerance distribution, F, on the estimator based on G, the asymptotic efficiency, $E_{aA}(G \mid F)$, is computed

$$E_{aA}(G \mid F) = \frac{MSE_{aA}(\mu^* \mid G)}{MSE_{aA}(\mu^* \mid F)} .$$

Tables 9.1.a, 9.1.b and 9.1.c contain computational results for F normal, logistic and uniform when the model is angular. The results indicate that the differences due to the several tolerance distributions are negligible even for N of 100, when the tolerance distributions are equated on distance between the 20^{th} and 80^{th} percentiles and the expected proportion responding is not too far from 50 percent, say between 15 percent and 85 percent. However, when the distributions are equated on standard deviation the asymptotic mean square error does show a marked decrease due to bias for N of 100 for some values of P (e.g. for the logistic, E=.66

Table 9.1.a

Asymptotic Efficiency (E_{aA}) for an RBAN Angular Estimator for Three Tolerance Distributions having Standard Deviations the Same as the Model (One Dose Level with N Subjects)

Expected Per Cent Responding at the Single Dose Level	$E_{aA}(G \mid F)$ 1/					
	N = 10			N = 100		
	Logistic (F)	Normal (F)	Uniform (F)	Logistic (F)	Normal (F)	Uniform (F)
1	.43	.66	.69	.07	.17	.18
3	.92	.96	.89	.56	.71	.45
5	1.00	1.00	.97	.99	1.00	.78
7	.96	.99	1.00	.74	.93	.97
10	.91	.97	1.00	.51	.79	.96
15	.87	.96	.97	.40	.69	.78
20	.87	.95	.96	.34	.68	.69
25	.89	.95	.96	.45	.72	.68
30	.92	.97	.96	.54	.78	.72
35	.95	.98	.97	.66	.85	.79
40	.98	.99	.99	.81	.93	.88
45	.99	1.00	1.00	.94	.98	.97
50	1.00	1.00	1.00	1.00	1.00	1.00

1/ $E_{aA}(G \mid F) = \dfrac{MSE_{aA}(\mu_1^* \mid G)}{MSE_{aA}(\mu_1^* \mid F)}$

where F is the true tolerance distribution and G is the angular model. See (9.1) for the definition of μ_1^*.

Table 9.1.b

Asymptotic Efficiency (E_{aA}) for an RBAN Angular Estimator for Three Tolerance Distributions having Distances Between the 20th and 80th Percentiles the Same as the Model (One Dose Level with N Subjects)

Expected Per Cent Responding at the Single Dose Level	$E_{aA}(G \mid F)$ 1/					
	N = 10			N = 100		
	Logistic (F)	Normal (F)	Uniform (F)	Logistic (F)	Normal (F)	Uniform (F)
1	.15	.37	.50	.02	.06	.09
3	.40	.68	.68	.06	.18	.18
5	.62	.84	.81	.14	.34	.30
7	.78	.92	.88	.27	.53	.44
10	.92	.97	.95	.53	.78	.68
15	.99	1.00	.99	.92	.97	.94
20	1.00	1.00	1.00	1.00	1.00	1.00
25	.99	1.00	11.00	.98	.99	.99
30	.99	1.00	1.00	.96	.99	.97
35	1.00	1.00	1.00	.97	1.00	.97
40	1.00	1.00	1.00	.99	1.00	.99
45	1.00	1.00	1.00	.99	1.00	1.00
50	1.00	1.00	1.00	1.00	1.00	1.00

1/ $$E_{aA}(G \mid F) = \frac{MSE_{aA}(\mu_1^* \mid G)}{MSE_{aA}(\mu_1^* \mid F)}\ .$$

where F is the true tolerance distribution and G is the angular model. See (9.1) for the definition of μ_1^*.

Table 9.1.c

Asymptotic Efficiency (E_{aA}) for an RBAN Angular Estimator for Three Tolerance Distributions having Information the Same as the Model (One Dose Level with N Subjects)

Expected Per Cent Responding at the Single Dose Level	$E_{aA}(G \mid F)$ 1/					
	N = 10			N = 100		
	Logistic (F)	Normal (F)	Uniform (F)	Logistic (F)	Normal (F)	Uniform (F)
1	.11	.01	.01	.01	.00	.00
3	.20	.04	.04	.02	.00	.00
5	.30	.09	.10	.04	.01	.01
7	.50	.17	.19	.07	.02	.02
10	.58	.32	.38	.12	.04	.06
15	.64	.64	.70	.15	.15	.19
20	.80	.71	.75	.28	.20	.23
25	.89	.86	.88	.45	.38	.42
30	.95	.94	.95	.66	.60	.64
35	1.00	1.00	1.00	1.00	1.00	1.00
40	1.00	1.00	1.00	1.00	1.00	1.00
45	1.00	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	1.00	1.00

1/ $$E_{aA}(G \mid F) = \frac{MSE_{aA}(\mu_1^* \mid G)}{MSE_{aA}(\mu_1^* \mid F)}$$

where F is the true tolerance distribution and G is the angular model. See (9.1) for the definition of μ_1^*.

at P=35 percent, and E=.34 at P=20 percent). Equating information shows even larger effects on the efficiency for small values of P.

The results for the angular model and estimator have been corroborated by repeating the computations for the logistic model and its estimator. Results analogous to those in Table 9.1.b are presented in Table 9.2 for the logistic estimator.

Since the results presented are asymptotic approximations, it is of interest to see whether the relationships indicated by these computations are valid in the range of sample size used. Exact computations analogous to Table 9.1.c are presented in Table 9.3 for N=10.

9.3 Two Level Experiment

Again let the model and estimator be based on the angular distribution (Table 2.2). Let the experiment consist of $\frac{1}{2}N$ subjects tested at $x=x_1= -\frac{1}{2}$ and $\frac{1}{2}N$ subjects tested at $x=x_2=\frac{1}{2}$. Let p_i and P_i be the observed and expected proportions respectively, i=1,2. Let $y_i=\sin^{-1}\sqrt{p_i} - \pi/4$.

If the scale parameter, β, is known, the RBAN estimator used will be denoted by μ_2^* and is:

$$\mu_2^* = -\bar{y}/\beta \qquad (9.5)$$

where $\bar{y} = \frac{1}{2}(y_1+y_2)$

Table 9.2

Asymptotic Efficiency (E_{aA}) of an RBAN Logistic Estimator for Three Tolerance Distributions having Distances Between the 20th and the 80th Percentiles the Same as the Model (One Dose Level with N Subjects)

Expected Per Cent Responding at the Single Dose Level	$E_{aA}(G \mid F)$ 1/					
	N = 10			N = 100		
	Normal (F)	Angular (F)	Uniform (F)	Normal (F)	Angular (F)	Uniform (F)
1	.94	.79	.65	.63	.27	.16
3	.96	.83	.67	.70	.32	.17
5	.97	.88	.74	.79	.42	.22
7	.99	.92	.81	.86	.55	.30
10	.99	.96	.90	.93	.74	.48
15	1.00	1.00	.98	.99	.94	.85
20	1.00	1.00	1.00	1.00	1.00	1.00
25	1.00	1.00	1.00	1.00	.98	.94
30	1.00	1.00	1.00	1.00	.97	.89
35	1.00	1.00	1.00	1.00	.97	.88
40	1.00	1.00	1.00	1.00	.98	.93
45	1.00	1.00	1.00	1.00	.99	.98
50	1.00	1.00	1.00	1.00	1.00	1.00

1/ $$E_{aA}(G \mid F) = \frac{MSE_{aA}(\mu_1^* \mid G)}{MSE_{aA}(\mu_1^* \mid F)}$$

where F is the true tolerance distribution and G is the logistic model.

Table 9.3

Exact Efficiency (E_a) for an RBAN Angular Estimator for Three Tolerance Distributions having Information Equal to that for the Model (One Dose Level with 10 Subjects)

Expected Per Cent Responding at the Single Dose Level	$E_a(G \mid F)$ [1/]		
	Logistic (F)	Normal (F)	Uniform (F)
1	.06	.01	.01
3	.17	.06	.06
5	.29	.18	.21
7	.41	.37	.44
10	.56	.66	.76
15	.76	.94	.98
20	.89	1.01	1.02
25	.95	1.01	1.02
30	.98	1.01	1.01
35	.99	1.00	1.00
40	1.00	1.00	1.00
45	1.00	1.00	1.00
50	1.00	1.00	1.00

1/ $E_a(G \mid F) = \dfrac{MSE_a(\mu_1^* \mid G)}{MSE_a(\mu_1^* \mid F)}$

where F is the true tolerance distribution and G is the angular model.

If the scale parameter is not known, the RBAN estimator used will be denoted by $\mu_2^{*'}$ and is:

$$\mu_2^{*'} = -\frac{\bar{y}}{y_2 - y_1} \tag{9.6}$$

The asymptotic means, variances and mean square errors for the two estimators, for given P_1 and P_2 are:

$$E_{aA}(\mu_2^*) = -\frac{Y_1 + Y_2}{2\beta} \tag{9.7}$$

where $Y_i = \sin^{-1}\sqrt{P_i} - \pi/4$

$$V_{aA}(\mu_2^*) = \frac{1}{4N\beta^2} \tag{9.8}$$

$$MSE_{aA}(\mu_2^*) = \frac{1}{4N\beta^2} + (\mu + \frac{Y_1 + Y_2}{2\beta})^2 \tag{9.9}$$

$$E_{aA}(\mu_2^{*'}) = -\frac{Y_1 + Y_2}{2(Y_2 - Y_1)} \tag{9.10}$$

$$V_{aA}(\mu_2^{*'}) = \frac{1}{2N} \frac{Y_2^2 + Y_1^2}{(Y_2 - Y_1)^4} \tag{9.11}$$

$$MSE_{aA}(\mu_2^{*'}) = \frac{1}{2N} \frac{Y_2^2 + Y_1^2}{(Y_2 - Y_1)^4} + \left[\mu + \frac{Y_2 + Y_1}{2(Y_2 - Y_1)}\right]^2 \tag{9.12}$$

If the tolerance distribution is angular, $\mu_2^{*'}$ is consistent; and if, in addition, the scale parameter is β then μ_2^* is consistent.

Denote the angular tolerance distribution by G. If the true tolerance distribution is $F \neq G$, then both estimators will be inconsistent. The asymptotic mean square error can be determined for given F and N, for specified values of P_1 and P_2. Tables 9.4.a and 9.4.b present asymptotic efficiencies of μ_2^* and $\mu_2^{*\prime}$ for F taken to be logistic when the model is angular. The results in Tables 9.4.a and 9.4.b show that in both cases the effect of a change in F from angular to logistic has little effect on the asymptotic mean square error of the angular estimator.

9.4 The 2k+1 Level Experiment

Again let the model and estimator be based on the angular distribution. Let the experiment consist of $\frac{N}{2k+1}$ subjects tested at each of 2k+1 levels. Let the dose levels be

$$-\frac{1}{2}, \frac{-k+1}{2k}, \ldots -\frac{1}{2k}, 0, \frac{1}{2k}, \ldots \frac{k-1}{2k}, \frac{1}{2}.$$

Let the observed proportions be denoted by p_i and the expected proportions by P_i. Let $y_i = \sin^{-1}\sqrt{p_i} - \pi/4$.

If the scale parameter, β, is known, then the maximum likelihood estimator is denoted by μ_5^* and is

$$\mu_5^* = -\frac{\bar{y}}{\beta} \tag{9.13}$$

$$\text{where } \bar{y} = \frac{\sum_{-k}^{k} y_i}{2k+1}$$

The asymptotic mean, variance and mean square error of the

Table 9.4.a

Asymptotic Efficiency (E_{aA}) for an RBAN Angular Estimator for the Logistic Tolerance Distribution (Scale Parameter Known: Two Dose Levels, $\frac{1}{2}N$ Subjects at each Dose Level)

Expected Per Cent Responding		$E_{aA}(G \mid F)$ 1/	
Lower Level (P_1)	Upper Level (P_2)	N = 10	N = 100
10	15	.70	.19
10	20	.80	.28
10	30	.93	.56
10	40	.98	.88
10	50	1.00	1.00
10	60	.99	.94
10	70	.98	.88
10	80	.98	.89
10	90	1.00	1.00
20	40	1.00	.97
20	60	1.00	.99
20	80	1.00	1.00
40	50	1.00	1.00
40	60	1.00	1.00

1/ $$E_{aA}(G \mid F) = \frac{MSE_{aA}(\mu_2^* \mid G)}{MSE_{aA}(\mu_2^* \mid F)}$$

where F is the logistic tolerance distribution and G is the angular model. See (9.5) for the definition of μ_2^*.

Table 9.4.b

Asymptotic Efficiency (E_{aA}) for an RBAN Angular Estimator for the Logistic Tolerance Distribution (Scale Parameter Unknown: Two Dose Levels, $\frac{1}{2}$N Subjects at each Dose Level)

Expected Per Cent Responding Lower Level (P_1)	Upper Level (P_2)	$E_{aA}(G \mid F)$ 1/ N = 10	N = 100
10	15	1.00	.96
10	20	.99	.92
10	30	.99	.91
10	40	1.00	.96
10	50	1.00	1.00
10	60	.99	.96
10	70	.98	.89
10	80	.98	.89
10	90	1.00	1.00
20	40	1.00	.99
20	60	1.00	1.00
20	80	1.00	1.00
40	50	1.00	1.00
40	60	1.00	1.00

1/ $$E_{aA}(G \mid F) = \frac{MSE_{aA}(\mu_2^{*\prime} \mid G)}{MSE_{aA}(\mu_2^{*\prime} \mid F)}$$

where F is the logistic tolerance distribution and G is the angular model. See (9.6) for the definition of $\mu_2^{*\prime}$.

estimator are, for given P_i:

$$E_{aA}(\mu_5^*) = -\frac{\bar{Y}}{\beta} \tag{9.14}$$

where $Y_i = \sin^{-1}\sqrt{P_i} - \pi/4$

and $\bar{Y} = \dfrac{\sum_{-k}^{k} Y_i}{2k+1}$

$$V_{aA}(\mu_5^*) = \frac{1}{4N\beta^2} \tag{9.15}$$

$$MSE_{aA}(\mu_5^*) = \frac{1}{4N\beta^2} + (\mu + \frac{\bar{Y}}{\beta})^2 \tag{9.16}$$

When the distribution is angular, the estimator is consistent. Let the angular tolerance distribution be denoted by G. For $F \neq G$, the asymptotic mean square error can be calculated. For G located to give a specified set of P_i a comparable F must be chosen. An F cannot be chosen which will give the same values as the model at all levels. In the computations F was chosen to give the same values at the end-points of the range of dose levels, i.e. at x_{-k} and x_k. Table 9.5 presents the results of computations for the case of five dose levels, angular estimator, and the logistic tolerance distribution. The effect of the F on the asymptotic mean square error is again seen to be negligible. In fact the results for five levels, β known, duplicate almost exactly the results for two levels, β known (Table 9.4.a).

Table 9.5

Asymptotic Efficiency (E_{aA}) of an RBAN Angular Estimator for the Logistic Tolerance Distribution (Scale Parameter Known: Five Dose Levels, N/5 Subjects at each Dose Level)

Expected Per Cent Responding Lowest Dose (P_1)	Highest Dose (P_5)	$E_{aA}(G \mid F)$ 1/ N = 10	N = 100
10	15	.70	.19
10	20	.79	.27
10	30	.91	.50
10	40	.96	.74
10	50	.99	.92
10	60	1.00	.99
10	70	1.00	1.00
10	80	1.00	1.00
10	90	1.00	1.00

1/ $$E_{aA}(G \mid F) = \frac{MSE_{aA}(\mu_5^* \mid G)}{MSE_{aA}(\mu_5^* \mid F)}$$

where F is the logistic tolerance distribution and G is the angular model. See (9.14) for the definition of μ_5^* .

9.5 Summary of Wrong Model Investigation

The above computations indicate that in the case of parametric estimation a model can be chosen and slight deviations of the true tolerance distribution from the assumed functional form will have little effect on the mean square error of the estimator. (It is important that the tolerance distribution have about the same spread, measured in interpercentile deviation, anticipated in the model, if the scale parameter is assumed known.)

Previous sections (7, 8) have shown that the Spearman estimator is a very efficient estimator compared with the amount of information available in the experiment. This section indicates that this efficiency would not be greatly increased if comparisons were based on the mean square error of fully efficient parametric estimators taking into consideration the possibility of using a wrong model. In spite of this apparent robustness of the parametric estimators, the Spearman estimator is recommended for use in most quantal assay experiments because of its simplicity and high efficiency.

REFERENCES

(1) Adams, E., Smithsonian Mathematical Formulae and Tables of Elliptic Functions, Washington, D.C.: Smithsonian Institution (1939).

(2) Armitage, P. and Allen, I., "Methods of estimating the L.D. 50 in quantal response data," Journal of Hygiene, 48 (1950) 298-322.

(3) Berkson, Joseph and Dews, Peter B., "On the Error of Bioassay with Quantal Response," Statistics and Mathematics in Biology, Ames: The Iowa State College Press (1954) 361-370.

(4) Berkson, Joseph, "A statistically precise and relatively simple method of estimating the bioassay with quantal response, based on the logistic function," Journal of the American Statistical Association, 48 (1955) 565-599.

(5) Berkson, Joseph, "Maximum likelihood and minimum χ^2 estimates of the logistic function," Journal of the American Statistical Association, 50 (1955) 130-162.

(6) Berkson, Joseph, "Tables for use in estimating the normal distribution function by normit analysis. Part I. Description and use of the tables. Part II. Comparison between minimum normit χ^2 estimates and the maximum likelihood estimator," Biometrika, 44 (1957) 411-435.

(7) Blackith, R.E., "Bioassay systems for the pyrethrins III. Application of the twin cross-over design to crawling insect assays," Annals of Applied Biology, 37 (1950) 508-515.

(8) Bliss, C. I., The Statistics of Bioassay, New York: Academic Press (1952).

(9) Bross, Irwin, "Estimates of the L.D.$_{50}$: A critique," Biometrics, 6 (1950) 413-423.

(10) Cornfield, Jerome and Mantel, Nathan, "Some new aspects of the application of maximum likelihood to the calculation of the dosage response curve," Journal of the American Statistical Association, 45 (1950) 181-210.

(11) Cramer, Harald, Mathematical Methods of Statistics, Princeton: Princeton University Press (1951).

(12) Doob, J.L., Stochastic Processes, New York: John L. Wiley and Sons, Inc. (1953).

(13) Emmens, C., Principles of Biological Assay, London: Chapman and Hall (1948).

(14) Epstein, Benjamin and Churchman, C. West, "On the statistics of sensitivity data," Annals of Mathematical Statistics, 15 (1944) 90-96.

(15) Finney, D. J., "The estimation of the mean of a normal tolerance distribution," Sankyā, 10 (1950) 341-360.

(16) Finney, D. J., "The estimation of the E.D. 50 for a logistic response curve," Sankyā, 12 (1952) 121-136.

(17) Finney, D.J., Statistical Method in Biological Assay, New York: Hafner Publishing Company (1952) section 20.11.

(18) Finney, D.J., Probit Analysis (second edition), London: Cambridge University Press (1952).

(19) Gaddum, J. H., "Reports on Biological Standards. III. Methods of biological assay depending on a quantal response," Medical Research Council, Special Report Series, (1933) 183.

(20) Gyorgy, P., Vitamin Methods, volume 2, New York: Academic Press Inc. (1951).

(21) Haley, David C., "Estimation of the dosage-mortality relationship when the dose is subject to error," Technical Report No. 15, Applied Mathematics and Statistics Laboratory, Stanford University (1952).

(22) Irwin, J. O., "Statistical method applied to biological assays," Journal of the Royal Statistical Society, Suppl.4 (1937) 1-60.

(23) Karber, G., "Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche," Archiv fur Experimentelle Pathologie and Pharmakologie, 162 (1931) 480-487.

(24) Kolmogorov, A. N., Foundations of the Theory of Probability, New York: Chelsea Publishing Company (1950).

(25) Meier, Paul, "Safety testing in poliomyelitis vaccine," Science, 125 (1957) 1067-1071.

(26) Neyman, Jerzy, "Contribution to the theory of the χ^2 test," Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, Berkeley: University of California Press (1949) 239-274.

(27) Pharmacopeia of the United States, fifteenth edition, Easton, Pa.: Mack Printing Company (1955).

(28) Spearman, C., "The method of 'right and wrong cases' ('constant stimuli') without Gauss' formulae," British Journal of Psychology, 2 (1908) 227-242.

(29) Taylor, William F., "Distance functions and regular best asymptotically normal estimates," Annals of Mathematical Statistics, 24 (1953) 85-92.

(30) Weaver, O.R., "Using attributes for measurement of continuous variables in the rocket industry," Proceedings, Twelfth Midwest Quality Control Conference, American Society for Quality Control (1957).

APPENDIX I. THE DISTRIBUTION FUNCTION AND INFORMATION FOR THE INFINITE EXPERIMENT

I.1 Distribution Function for the Infinite Experiment

Lemma I.1 Let $F(x)$ be a tolerance distribution. Specify a value for d. Choose x_o randomly from the interval $(0,d)$. Let $x_i = x_o + id$, $i = 0, \pm1, \pm2, \ldots$. Take n observations on a Bernoulli variable with expected value $F(x_i)$ for $i = 0, \pm1, \pm2, \ldots$. Let all the observations be independent. Let r_i be the number of responses at x_i. Then this infinite experiment determines a probability function for the r_i on the infinite sample space.

Proof: Since the r_i are independent sums of independent Bernoulli variables, the distribution functions for finite sets of the r_i satisfy the consistency conditions on page 29 of Kolmogorov (24). The lemma follows immediately from the theorem on the same page.

I.2 Information for the Infinite Experiment Scale Parameter Known

The information for the infinite experiment will be defined as the limit as $k \longrightarrow \infty$ of the information for the finite experiment, with dose levels $x_i = x_o + id$, $i = 0, \pm1, \pm2, \ldots \pm k$, and x_o randomly chosen.

The information for the finite experiment with fixed x_o is

$$I_k(x_o) = E\left[\frac{\partial \ln h}{\partial \mu}\right]^2$$

where h is the density of the finite sequence of binomial variables r_i, $i=0, \pm 1, \pm 2, \ldots \pm k$.

$$h = \prod_{-k}^{k} \binom{n}{r_i} F_i^{r_i} (1-F_i)^{n-r_i}$$

$$I_k(x_o) = \sum_{-k}^{k} \frac{nF_{\mu i}^2}{F_i(1-F_i)} \qquad \text{where } F_{\mu i} = \frac{\partial F}{\partial \mu}\Big|_{x=x_i}$$

The information for the infinite experiment, I, will be defined as

$$I = \lim_{k \to \infty} E_{x_o}\left[I_k(x_o)\right]$$

$$I = \lim_{k \to \infty} \int_o^d \frac{1}{d} \sum_{-k}^{k} \frac{nF_{\mu i}^2}{F_i(1-F_i)} \, dx_o$$

$$I = \frac{n}{d} \int_{-\infty}^{\infty} \frac{F_\mu^2}{F(1-F)} \, dt$$

It can be shown that if μ is a translation parameter then I is finite.

I.3 Information for the Infinite Experiment, Scale Parameter Unknown

Let F be written $F\left[\beta(x-\mu)\right]$. The information matrix for the infinite experiment will be defined as the matrix obtained as $k \to \infty$ for the finite experiment.

The information matrix (I_{ij}) for the finite experiment with given x_o is

$$\left[I_{ij}(x_o)\right] = \begin{bmatrix} \sum_{-k}^{k} \dfrac{nF_{\mu i}^2}{F_i(1-F_i)} & \sum_{-k}^{k} \dfrac{nF_{\mu i}F_{\beta i}}{F_i(1-F_i)} \\ \sum_{-k}^{k} \dfrac{nF_{\mu i}F_{\beta i}}{F_i(1-F_i)} & \sum_{-k}^{k} \dfrac{nF_{\beta i}^2}{F_i(1-F_i)} \end{bmatrix}$$

If x_o is randomly chosen and the limit is taken as $k \longrightarrow \infty$ the matrix is:

$$(I_{ij}) = \begin{bmatrix} \dfrac{\beta n}{d}\displaystyle\int_{-\infty}^{\infty} \dfrac{F'^2}{F(1-F)}\,dt & -\dfrac{n}{\beta d}\displaystyle\int_{-\infty}^{\infty} \dfrac{tF'^2}{F(1-F)}\,dt \\ -\dfrac{n}{\beta d}\displaystyle\int_{-\infty}^{\infty} \dfrac{tF'^2}{F(1-F)}\,dt & \dfrac{n}{\beta^3 d}\displaystyle\int_{-\infty}^{\infty} \dfrac{t^2F'^2}{F(1-F)}\,dt \end{bmatrix}$$

The element of the inverse matrix corresponding to μ is:

$$I^{11} = \frac{\dfrac{n}{\beta^3 d}\displaystyle\int \dfrac{t^2F'^2}{F(1-F)}\,dt}{\dfrac{n^2}{d^2\beta^2}\displaystyle\int \dfrac{t^2F'^2}{F(1-F)}\,dt \int \dfrac{F'^2}{F(1-F)}\,dt - \dfrac{n^2}{d^2\beta^2}\left[\displaystyle\int \dfrac{tF'^2}{F(1-F)}\,dt\right]^2}$$

$$= \frac{1}{\dfrac{n}{\beta d}\displaystyle\int \dfrac{F'^2}{F(1-F)}\,dt} \quad \frac{A_1^2 + A_2}{A_2}$$

where

$$A_1 = \frac{\int t \frac{F'^2}{F(1-F)}\, dt}{\int \frac{F'^2}{F(1-F)}\, dt}$$

$$A_2 = \frac{\int (t-A_1)^2 \frac{F'^2}{F(1-F)}\, dt}{\int \frac{F'^2}{F(1-F)}\, dt}$$

APPENDIX II. BERNOULLI PERIODIC FUNCTIONS AND THE EULER-MACLAURIN FORMULAE

II.1 The Bernoulli Periodic Functions

The Bernoulli periodic functions, $P_n(t)$, and the Euler-MacLaurin formulae are presented here. For details of the development of the formulae the reader is referred to Cramer (11) pp. 122-125.

The Bernoulli periodic functions are:

$$P_{2n}(t) = \sum_{k=1}^{\infty} \frac{\cos k\pi t}{2^{2n-1}(k\pi)^{2n}} \qquad n=1,2,\ldots$$

$$P_{2n+1}(t) = \sum_{k=1}^{\infty} \frac{\sin k\pi t}{2^{2n}(k\pi)^{2n+1}} \qquad n=0,\ 1,\ 2,\ldots$$

These functions satisfy

$$P_n'(t) = (-1)^{n-1}\, P_{n-1}(t)\ .$$

The first three Bernoulli functions are (see reference 1, p. 138):

$$P_1(t) = \frac{1}{2} - t \qquad 0 < t < 1$$

$$P_2(t) = \frac{1}{12} - \frac{t}{2} + \frac{t^2}{2} \qquad 0 < t < 1$$

$$P_3(t) = \frac{t}{12} - \frac{t^2}{4} + \frac{t^3}{6} \qquad 0 < t < 1$$

In paragraph 4.2.7 the $\sup_t P_n(t)$ is desired for n=1,2 and 3.

Using the polynomial expressions, the following are obtained:

$$\sup_t P_1(t) = \frac{1}{2}$$

$$\sup_t P_2(t) = \frac{1}{12}$$

$$\sup_t P_3(t) = .0080\ldots$$

II.2 The Euler-MacLaurin Formulae

Let x_o and d be constants and let the i^{th} term in a finite sum be $g(x_o+id)$. Let g be continuous with a continuous derivative g'. Then the first Euler-MacLaurin formula is:

$$\sum_{-k}^{k} g(x_o+id) = \int_{-k}^{k} g(x_o+xd)dx + \tfrac{1}{2}g(x_o-kd) + \tfrac{1}{2}g(x_o+kd)$$
$$-d\int_{-k}^{k} P_1(x)g'(x_o+xd)dx$$

If g has continuous derivatives of higher orders repeated integration by parts gives the Euler-MacLaurin formulae:

$$\sum_{-k}^{k} g(x_o+id) = \int_{-k}^{k} g(x_o+id)dx + \frac{1}{2}\, g(x_o-kd) + \frac{1}{2}\, g(x_o+kd)$$
$$-\sum_{1}^{s} \frac{B_{2i}}{(2i)!}\, d^{2i-1}\left[g^{(2i-1)}(x_o-kd)-g^{(2i-1)}(x_o+kd)\right]$$
$$+(-1)^{s+1} d^{2s+1}\int_{-k}^{k} P_{2s+1}(x)g^{(2s+1)}(x_o+xd)dx$$

where s is any non-negative integer, g^j must exist for $j=1,2,\ldots(2s+1)$, and B_j are the Bernoulli numbers, defined by:

$$\frac{x}{e^x-1} = \sum_{o}^{\infty} \frac{B_j}{j!}\, x^j$$

The first Euler-MacLaurin formula given above can be applied to an infinite sum to obtain:

$$\sum_{-\infty}^{\infty} g(x_o+id) = \int_{-\infty}^{\infty} g(x_o+xd)dx - d\int_{-\infty}^{\infty} P_1(x)g'(x_o+xd)dx$$

provided that the series and the two integrals converge.

APPENDIX III. ESTIMATION OF THE VARIANCE OF THE SPEARMAN ESTIMATOR

III.1 An estimator for $V(\bar{x})$

An obvious estimator for the variance of the Spearman estimator is:

$$s_{\bar{x}}^2 = \frac{d^2}{n-1} \sum_{-\infty}^{\infty} p_i q_i \qquad (\text{if } s_{\bar{x}}^2 \text{ converges}).$$

Then

$$E(s_{\bar{x}}^2 \mid x_o) = V(\bar{x} \mid x_o)$$

The following results can be obtained for this estimator for the infinite experiment with random choice of x_o, following the same methods used for obtaining the characteristics of $\bar{x}$.

If F has a first moment, $s_{\bar{x}}^2$ converges with probability one and has the following mean and mean square error.

$$E(s_{\bar{x}}^2) = \frac{d}{n} \int F(1-F)dx = V(\bar{x})$$

$$MSE(s_{\bar{x}}^2) = \left(\frac{d}{n}\right)^3 \left[\frac{n^2-6n+6}{n(n-1)}\right] \int F^2(1-F)^2 dx$$

$$+ \frac{1}{n} \left[\int F(1-F)dx\right] + E_{x_o}\left[B^2(s_{\bar{x}}^2 \mid x_o)\right]$$

where $B^2(s_{\bar{x}}^2 \mid x_o)$ denotes the bias of $s_{\bar{x}}^2$ as an estimate of $V(\bar{x})$ (5.4), conditional on x_o.

Unlike the case of $\bar{x}$, where it can be shown that the $MSE(\bar{x})$ is a minimum for n=1 when n'=n/d is fixed, for $s_{\bar{x}}^2$

the optimum choice of n for n' fixed will depend on F. For the normal tolerance distribution $MSE(s\frac{2}{x})$ decreases as n increases but the decrease is negligible for n greater than 4. In this case, then, the optimum designs for estimation of μ and of $\hat{V}(\bar{x})$ do not agree, but a good compromise would be to choose $n=4$, say.

III.2 <u>An Alternative Estimator for $\hat{V}(\bar{x})$ Based on the Second Moment of F</u>

Table 5.1 shows $\hat{V}(\bar{x})$ to be a function of the scale parameter:

$$\hat{V}(\bar{x}) = \frac{d\sigma}{n} C_F \qquad \text{(III.1)}$$

where C_F is a constant depending on the parametric formulation of the tolerance distribution. The constants for the normal, logistic, angular and uniform distributions are .5642, .5513, .5750 and .5774 respectively.

If the constant C_F is considered over all distribution functions, the function defined by:

$$P(x = -\frac{1}{a}) = \frac{a^2}{1+a^2} \qquad a > 0$$

$$P(x = a) = \frac{1}{1+a^2}$$

will have $C_F = \frac{a}{1+a^2}$, so that C_F is arbitrarily close to zero for small a. There is a unique maximizing function for C_F

over any given finite interval (see Rustagi[1]), and it can be shown by a calculus of variations argument that the uniform distribution is this function ($C_F = .5774$).

An alternative estimator ($s^{*2}_{\bar{x}}$) for $V(\bar{x})$ suggested by (III.1) is:

$$s^{*2}_{\bar{x}} = \frac{d}{n} C_F \sqrt{\sum_{-\infty}^{\infty} (x_i + {}^d/_2 - \bar{x})^2 \; (p_{i+1} - p_i)}$$

The estimate of σ used in $s^{*2}_{\bar{x}}$ is the Spearman type of estimator suggested by Epstein and Churchman (14) and shown by Cornfield and Mantel to be an algebraic approximation to the maximum likelihood estimator of σ for F logistic.

[1] See Rustagi, Jagdish Sharan, "On Minimizing and Maximizing a Certain Integral with Statistical Implications," Annals of Mathematical Statistics, 28 (1957) 309-328.

School of Business Administration
University of Minnesota

BASIC DISTRIBUTION LIST FOR UNCLASSIFIED TECHNICAL REPORTS

for contract Nonr 2582(00), Task NR 042-200

ADDRESS	
Head, Statistics Branch Office of Naval Research Washington 25, D. C.	2
Commanding Officer Office of Naval Research Branch Office Navy #100 Fleet Post Office New York, New York	2
ASTIA Document Service Center Arlington Hall Station Arlington 12, Virginia	5
Office of Technical Services Department of Commerce Washington 25, D. C.	1
Technical Information Officer Naval Research Laboratory Washington 25, D. C.	6
Professor T. W. Anderson Department of Mathematical Statistics Columbia University New York 27, New York	1
Professor Z. W. Birnbaum Laboratory of Statistical Research Department of Mathematics University of Washington Seattle 5, Washington	1
Professor A. H. Bowker Applied Mathematics & Statistics Lab. Stanford University Stanford, California	1
Professor Ralph A. Bradley Department of Statistics & Statistical Laboratory Virginia Polytechnic Institute Blacksburg, Virginia	1
Professor Herman Chernoff Applied Mathematics & Statistics Lab. Stanford University Stanford, California	1

ADDRESS	
Professor W. G. Cochran Department of Statistics Harvard University Cambridge, Massachusetts	1
Professor Benjamin Epstein Applied Mathematics & Statistics Lab. Stanford University Stanford, California	1
Professor W. Hirsch Institute of Mathematical Sciences New York University New York 3, New York	1
Dr. Paul G. Hoel Department of Mathematics University of California Los Angeles 24, California	1
Professor Harold Hotelling Associate Director Institute of Statistics University of North Carolina Chapel Hill, North Carolina	1
Professor L. Hurwicz School of Business Administration University of Minnesota Minneapolis 14, Minnesota	1
Professor Leo Katz Department of Statistics Michigan State University East Lansing, Michigan	1
Professor Oscar Kempthorne Statistics Laboratory Iowa State College Ames, Iowa	1
Dr. Carl F. Kossack Statistical Laboratory Engineering Administration Building Purdue University Lafayette, Indiana	1

ADDRESS		ADDRESS	
Professor Gerald J. Lieberman Applied Mathematics and Statistics Laboratory Stanford University Stanford, California	1	Professor Evan J. Williams Institute of Statistics State College Section North Carolina State College Raleigh, North Carolina	1
Professor William G. Madow Department of Statistics Stanford University Stanford, California	1	Professor J. Wolfowitz Department of Mathematics Cornell University Ithaca, New York	1
Professor J. Neyman Department of Statistics University of California Berkeley 4, California	1		
Professor Herbert Robbins Mathematical Statistics Department Columbia University New York 27, New York	1		
Professor Murray Rosenblatt Department of Mathematics Indiana University Bloomington, Indiana	1		
Professor L. J. Savage Statistical Research Laboratory Chicago University Chicago 37, Illinois	1		
Dr. Herbert Solomon Teachers College Department of Statistics Columbia University New York 27, New York	1		
Professor Frank Spitzer Department of Mathematics University of Minnesota Minneapolis 14, Minnesota	1		
Professor S. S. Wilks Department of Mathematics Princeton University Princeton, New Jersey	1		